



MAURITIUS RESEARCH COUNCIL

A QUESTION ANSWER SYSTEM FOR THE MAURITIAN JUDICIARY

MAURITIUS RESEARCH COUNCIL

Address:

Level 6, Ebène Heights,
34, Cybercity,
Ebène 72201,
Mauritius.

Telephone: (230) 465 1235

Fax: (230) 465 1239

Email: mrc@intnet.mu

Website: www.mrc.org.mu

This report is based on work supported by the Mauritius Research Council under award number MRC/RUN-1301. Any opinions, findings, recommendations and conclusions expressed herein are the author's and do not necessarily reflect those of the Council.

PI-Name and Address

Assoc. Prof. (Dr) Rajendra Parsad Gunputh
Rm 1.15, 1st floor, New Academic Complex
Department of Law
Faculty of Law and Management
University of Mauritius
Reduit, Maka

MAURITIUS RESEARCH COUNCIL

FINAL REPORT

PART I- PROJECT IDENTIFICATION INFORMATION

1. Type the name of the MRC Scheme under which grant is made
Unsolicited Research and Innovation Grant Scheme
2. Award Dates (MM/YYYY) From: Nov 2015 To: July 2016
3. Organisation and Address
University of Mauritius
Reduit, Maka
4. Award Number: MRC/RUN-AO114/02
5. Project Title

A Question Answer System for the Mauritian Judiciary

Mauritius Research Council

MRC Grant Conditions require submission of a Final Report to the MRC no later than 30 days after the expiration date of the award. Final Reports for expired awards must be received before new awards can be made.

Use this format to provide a summary of the completed projects and technical information. Be sure to include your name and award number on each separate page. See below for more instructions.

PART II - SUMMARY OF COMPLETED PROJECT (for public use)

The summary (about 200 words) must be self-contained and intelligible to a scientifically or technically literate reader. Without restating the project title, it should begin with a topic sentence stating the project's major thesis. The summary should include, if pertinent to the project being described, the following items:

- The primary objectives and scope of the project
- The techniques or approaches used only to the degree necessary for comprehension
- The findings and implications stated as concisely and informatively as possible

PART III - TECHNICAL INFORMATION (for program management use)

List references to publications resulting from this award and describe in full primary data, samples, physical collections, inventions, software, etc., created or gathered in the course of the research and, if appropriate, how they are being made available to the research community. This section should be limited to 30 pages.

I certify to the best of my knowledge (1) the statement herein (excluding scientific hypotheses and scientific opinion) are true and complete, and (2) the text and graphics in this report as well as any accompanying publications or other documents, unless otherwise indicated, are the original work of the signatories or of individuals working under their supervision. I understand that willfully making a false statement or concealing a material fact in this report or any other communication submitted to MRC is a criminal offense.

Principal Investigator Signature:



A handwritten signature in black ink, appearing to read 'YV' followed by a flourish, is written over a diagonal line that crosses the signature box from the bottom-left to the top-right.

Date: \,.,,



A handwritten date '1/1/00' is written in black ink, with a diagonal line striking through it.

PART II - SUMMARY OF COMPLETED PROJECT

Law is a research-orientated profession and legal research is an activity that cost time and money. Information Technology is now revolutionising the way in which legal research is being done. In this work, we have implemented a web-based question answer and information retrieval system for the Mauritian Judiciary. On the *www.lawanswers.me* portal, users can enter their queries freely using natural language. The system processes the query by extracting the relevant keywords and discard those that do not carry much g information and then returns the relevant sections of law which contains these keywords or keyphrases. The system also returns a list of relevant Supreme Court cases. The user can decide on the number of results to be displayed. The user can also wish to have only the name of the relevant acts be displayed for certain keywords or keyphrases. The system does not require the user to know how the law is structured or how the knowledge-base is built in order to benefit from it. The portal can also be accessed via mobile devices without compromising any of its facilities or user-friendliness. It is hoped that the availability of information at the click of a button will help the human resources at the Mauritian Judiciary to become more efficient and this will contribute to the reduction of delays in the disposal of cases. As the old adage goes, "*justice delayed is justice denied*".

PART III - TECHNICAL INFORMATION

Table of Contents

I. Introduction	1
II. Objectives v/s Achievements	3
III. Question Answer Systems	4
i. Question Processing	5
ii. Document Processing	8
iii. Answer Extraction	8
IV. Related Works	9
V. An Overview of the Different Existing Question Answer Systems and Information Retrieval Systems	13
A. The Website of the Supreme Court of Mauritius	13
B. Lexis Nexis	17
C. Google v/s Google Scholar	21
D. General Purpose Question Answering System (Open-Domain)	22
E. Closed-Domain Question Answering Systems for the Medical Domain	25
F. Comparison of Answers by different Question Answering Systems	27
VI. Implementation	29
HTML5	29
CSS	29
PHP	29
MySQL	29
Example 1: Query in natural language	31
Example 2: Query in natural language	32
Example 3: Query as a bag of keywords	33
Example 4: Query as a keyphrase in double quotes	34
Example 5: Query as a keyphrase in double quotes with filter	35
Example 6: List of Acts	36
Example 7: Display of Acts	36
VII. Status of Outstanding Work from Progress Report Submitted on 08th March 2016	37
VIII. Conclusions	38
Publications	38
Software	38
Acknowledgements	38
REFERENCES	39

I. Introduction

The amount of available information in the legal sphere of Mauritius is constantly rising. Evidence from the Supreme Court's Annual Report 2014 depicts that the number of cases lodged before the Mauritian Judiciary has widely increased from 126, 645 in 2011 to 133, 423 in 2014 [1]. This has consequently also resulted in a rise in the number of outstanding cases from 39,336 in 2011 to 46,055 in 2014. On the other hand, the number of new laws that are being promulgated are also increasing at a much more rapid pace. A substantial number of amendments are also being made to the existing statutes every year.

It is commonly reported that access to legal information is a key to administer justice rightly in any democratic country [2]. However, we also live with the reality that legal professionals and actors, while operating in a highly intensive information environment often encounter considerable barriers in their quest to find relevant information or specific answers to their questions [3]. Lawyers, attorneys, legal professionals and other actors operating in the legal domain are often on the run for legal information search because they need to be constantly updated about the recent developments in the law, about legal precedents or even about how justice rendered in a particular case. The legal profession is particularly characterised by information and case law abundance and most importantly by a continuous flooding of amendments and additions to the legal statutes [4]. However, in most countries, there is a lack of all-inclusive or complete law collection and there are in most cases substantial gaps which require attention and redress. It is also unfortunate to note that unlike other fields, like in the medical field for example, that equally have tremendous amount of information available, the legal domain to our knowledge has barely any kind of question answering systems that would enable actors operating in the legal profession to ask a question in natural language and be able to get a specific answer to their queries.

Mauritius' position with respect to access to legal information is also quite restricted. With the rise of new emerging sectors like the ICT / BPO and the Ocean Economy, there will be also a parallel rise in the number of upcoming legislations. With all these new developments, it is progressively more and more difficult for actors operating in the legal arena to deliver quality work on time. Legal research constitutes a very large portion of the time that legal practitioners spend in preparing their cases. Information is required in all aspects for a legal practitioner to effectively deliver to his/her duties. Be it for giving a legal advice or drafting up a document or even to assist clients in Court, information is of utmost importance. With these issues in mind and considering the fact that there is a lack of all-inclusive platform that would effectively help the legal professionals or the public in general to obtain precise answers to their questions, we propose to develop an automatic question answering system to mainly assist and support legal practitioners in their day-to-day work.

This report proceeds as follows. In the next section, we list the main objectives of the proposed system, the achievements, the issues that have been addressed and the scope of our work. In section III, we give a detailed description of the architecture of question answer systems. Section IV gives an overview of related academic works that have been carried out in the field of question answering. We thereafter give an overview of the various existing question answer systems and information retrieval systems in Section V. Section VI provides details about the implementation of the web-based system. Section VII describes how the outstanding work after the submission of the progress report has been tackled and Section VIII concludes the report.

II. Objectives v/s Achievements

The main objectives of the proposed system are as follows:

- To create a database of statutes of the Republic of Mauritius;
- To create a database of past cases;
- To create a database of other legal materials such as statistical reports;
- To develop a software, using an appropriate programming language, to allow a user to ask a question and gets a short and concise answer within a few seconds;
- To evaluate the answers produced by the system;
- To identify potential weaknesses of the system and to correct them;
- To disseminate the appropriate knowledge to relevant stakeholders.

The following tasks have been completed between 27 November 2015 and 27 July 2016.

1. We have downloaded about 600 statutes from the website of the Supreme Court of Mauritius. Some of these documents have been pre-processed so that they can be entered into the knowledge base of our question answer system. The documents had to be converted to text and then split into appropriate sections.
2. We have also downloaded about 500 cases from the website of the Supreme Court of Mauritius. These have been categorised into their appropriate area of law such as constitutional law, criminal law, labour law, company law, etc. This categorisation is important for the Question Answer system in order to increase the relevancy of the results that have been provided.
3. Only acts and judgements have been considered so far in this work. Other legal materials may be incorporated in a future work.
4. A web-based Question Answer system has been implemented. The user can ask a query in natural language and receives a relevant reply almost instantaneously.
5. Some LLB and LPVC (Law Practitioner Vocational Course) students of the University of Mauritius had been asked to evaluate the system and their comments were taken into consideration in order to improve the system.
6. Some weaknesses were identified and corrected. For example, several users had requested that there should be a section from where all acts must be made available.
7. The outcomes from this work will be presented during the Research Week of the University of Mauritius, which is to be held in September 2016. We also hope to present this work in an international conference.

III. Question Answer Systems

A question answering (QA) system is basically a computer program that is designed to enable users to input a question in the system in a natural language and thereafter the system processes the question and outputs some results that precisely matches the question asked by the user. In other words, a QA system therefore accepts queries that are asked in a natural language or plain language, looks for the appropriate answers in different available document resources (or knowledge-base) and performs extraction and frames precise answers to the question asked [5][6].

More than ever before, we are nowadays living in an era where the quest for knowledge is deeply human. The various existing information systems and document retrieval systems bring with them a very large amount of documentation that makes it cumbersome for users or professionals to retrieve specific answers within limited time. Such systems, despite being very common tend to return a whole list of documents and relevant pages rather than returning specific answers [7]. QA is becoming a fast developing groundwork area that combines together result from Information Retrieval, Information Extraction and NLP [8] [9]. In order to return the best possible answers, QA systems query a structured database of information, also commonly known as a knowledge base. It may also formulate the result by combining information obtained from both local databases and from documents available of the web.

Question answering systems can be broadly classified as closed domain or open domain. As suggested by its name, the prior category of QA system would be one dealing with questions of determined domains that could range from health, life sciences, medicines, legal, automotive maintenance, among others. In most cases, closed domain QA systems have the ability to accept a restrictive category of questions and they tend to have the ability to respond to descriptive rather than procedural queries. On the other hand, open domain QA systems have the capacity of dealing with questions touching nearly all domains. Given their ability to depend on both general ontologies and world knowledge,

such systems would therefore be able to obtain the answers from greater amount of data [10].

In this section, emphasis is laid on the different components which underlie question answering systems [11] [12] [13]. There are three important phases, which form part of automatic QA systems, and these are mainly: Question Processing, Document Retrieval and thirdly, Answer Extraction [14] [15].

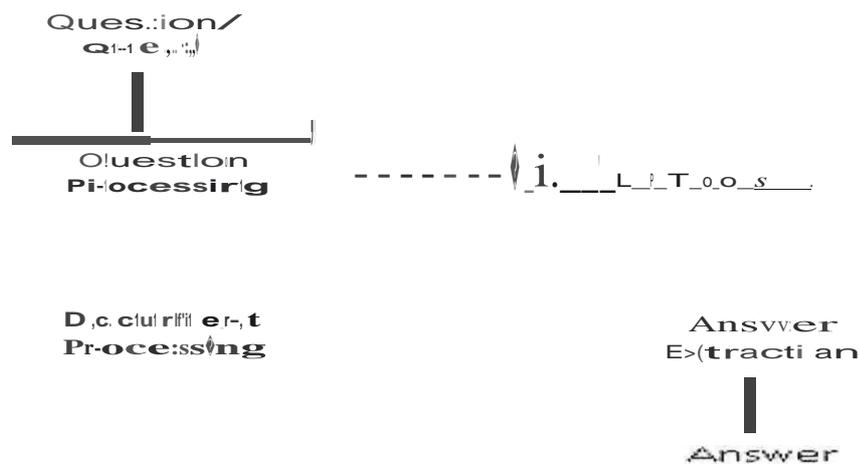


Figure 1: Architecture of a Question Answer System

i. Question Processing

The first module forming part of a QA system is question processing. It aims mainly at understanding the question asked by the user in a natural language. At this stage, the QA system will be analyzing the question asked in natural language and it will be using the different techniques of NLP such as Syntactic Analysis and Semantic Analysis. Python NLTK, Stanford NLP and CoreNLP are some of the tools which are generally used for this module [16]. Under Semantic Analysis, after receiving the user's question, some constraints must be imposed in order to limit the possibilities of getting immaterial answers to the questions asked. It would be mainly the retrieval engine which is responsible for handling the user's requests.

Question Classification

The classification of questions in different categories of sets is an important step in order to retrieve specific answers to various questions. The question posed in a natural language would normally contain the main information that would be required to get the answers. The type of answer to be answered would depend principally on the type of questions that have been asked. A tokenizer first breaks the question input into words; white spaces and stop words such as "and" or "the" are removed as they do not hold any importance in the question and thereafter, tokens are used for querying. Question classification will help in finding entity instanced in the various documents and would thereafter, choose the highest expected answer [18]. Focus identification, phrase or clause detection and frame detection are the various stages carried out by the QA system in the quest to fund the accurate answer. Focus identification helps in finding the category of answers looked for by the candidate. Question classification that can be used in question focus can be illustrated in the Table below:

Table 1. Type of Questions and Answers

Question Type	Answer Type
When	Time
Where	Place
Why	Reason
Describe	Description
Define	Definition
Who, Whom	Person
How old	Age
How much	Quantity
How long	Distance
What, which, name	Ambiguous

Query Formulation and Expansion

Query formulation and expansion is a set of methods applied to the input, which help to retrieve document in such a way that it improves the document retriever's performance [19] [20]. The focus should be put on how to expand queries. For example, "Who killed J.F.K", here the verb is fundamental for finding the correct answer. The text files used for document retrieval may or may not contain the word "to kill" though it contain the word "J.F.K", on the contrary the documents may contain verbs such as "to murder", "to assassinate", "to shoot". So a dictionary of such words should be maintained such as WordNet [21].

Query Expansion with WordNet

WordNet is a lexical database which is available online [21]. It is a dictionary, which contains noun, verb, adjectives and adverbs with their synonyms. WordNet allows expanding words in queries that have a semantic relationship.

Query Expansion with Stemming and Lemmatization

In creating a QA system and in NLP, further processing need to be effected so as to remove punctuations, understand complex vocabulary and to normalise linguistic variations in the documents. Stemming helps in expanding the query by providing word variants and derivations [22] [23]. Stemming is usually done using the Porter Stemmer [24] which is rule-based and it helps in finding the lexeme or the root of the word that remains unchanged and which is meaningful. Lemmatization however is a way of finding the lemma of conjugated words and help to find out the original word from which its plural or feminine word has been derived. Lemmatization helps in achieving accuracy in the query that is otherwise not possible with shortened or unclear stems.

ii. Document Processing

This module performs queries from the natural language question and depending on the question asked, the information retrieval engine searches for the different documents that could possibly contain the answers. The documents can be structured databases, text files, pdf documents, web pages or any other types of retrievable archives resources.

iii. Answer Extraction

In this module, different techniques are applied in order to retrieve the relevant answer and display it to the user. The three phases under answer extraction are: filtering, answer ranking and answer generation [17][25]. During filtering, answers which are highly irrelevant are discarded. The remaining answers are then ranked according to their weight (a score is calculate based on some factors) before finally outputting it to the user. In some systems, filtering is done after ranking. For example, only the top 3 or top 20% of the answers are kept and shown to the user.

IV. Related Works

The literature in the domain of question answering has been very rich and extensive. Several research papers have been written with respect to closed domain question answering systems under specific domains like medicine, tourism, music, life sciences, maintenance, etc. [26] [27] [28]. The aim of this part is to portray mainly the existing and related works that exist in the field of question answering, not necessarily limited to the legal domain.

Automatic question answering systems that existed earlier had as principal motive to provide rapid and natural access to expert knowledge. Its knowledge base was drawn principally from structured databases. The LUNAR system for example was evaluated by Woods *et al.* [29] as being capable of answering questions related to Apollo II moon rocks for the NASA Manned Spacecraft Centre.

Tahri and Tibermachine [30] used the DBpedia ontology as their knowledge base for their question answering system. A Support Machine Classifier also used for the purpose of question classification. Their research showed an accuracy of 68% which they claimed to be higher than the Qakis [31] and True Knowledge [32] systems. In the evaluation part of the system, two hundred questions were used.

A closed domain question answering system for the specific domain of tourism was proposed by Bhoir and Potey [33]. The system's knowledge base was mainly information gathered from the web and Natural Language Processing techniques are then used to process the information. Evaluation of the question answering system showed that it was able to give specific short answers or list to the questions asked in natural language. The various names of hotels, distances between regions, names of tourist attractions were some of the questions to which the system could generate precise answers.

Moussa and Kader (2011), from Port-Said University in Egypt, developed a powerful question answering system named QASYO [34]. QASYO accepts queries in natural

language and uses the YAGO Ontology to return answers [35]. The principal aim of developing such a system was to create a search experience that would not ask the user to know specific or technical vocabularies or to understand the system's knowledge base. The evaluation of the QASYO system had not really been tested in order to be in a position to assess its ability to answer precisely to queries asked nor are we in a position to comment on the user's experience. However, results show that the system was found to be highly superior compared to the internet-scale answering engine, True Knowledge [36]. Out of the 100 questions used to test and compare the performance of the QASYO system to that of the True Knowledge answering system, results demonstrated that the prior system could handle 91 questions correctly as compared to 40 that were handled by True Knowledge.

Derici *et al.* (2015) designed a question answer system in the Turkish language and was implemented for the closed domain of high school geography [37]. Their motive was to principally support their education. In terms of architecture, their system was very similar to that of the deep Question Answering technology used by Watson [38]. However, as they created their own data set for their knowledge base and created their own questions, it proved to be difficult to compare and assess their system in an objective manner.

Studies have also been conducted with respect to improving the relevance and preciseness of answers that question answering systems would generate. This is of high relevance to our creation of a question answering system for the Mauritian Judiciary as we ultimately have as objective to create a system that will be able to generate simple but precise answers at a time. With respect to improving the relevance of answers, Komiya *et al.* (2013) proved that query expansion, web relevance score and translation probability could significantly help in generating more precise answers [39]. Moschitti, however, found that considerable improvement can be achieved in a QA system if both the question and the answers are categorized [40]. Their experiments were based on the *Reuters-21578* dataset.

On the other hand however, Pakray et al. (2011) were of the view that using answer validation as part of their question answering process doubled the accuracy of correct answers obtained from 22.8% to 45.8% [41]. Moreover, an approach for selecting the best answer based on redundancy and the semantic relations between candidate answers was described by Mendes and Coheur [42]. Their technique showed a 10% improvement over the baseline experiment. Dong et al. (2011), in an attempt to widen the capability of LogAnswer which is a German question answering system and in an attempt to better adapt it to QA forums found that meta-classifier could be used in order to avoid producing incorrect and imprecise answers [43].

Studies carried out by Brill et al. [44] are somehow different from the studies explored earlier. They differ in that instead of considering methods of improving the correctness of answers returned by the system, they rather consider strategies to predict the likely situation where the system is going to return a wrong answer. They consider it as crucial for a QA system to know the likeliness of an answer to be a correct one in order to choose not to give an answer instead of giving a wrong answer. Therefore, in an attempt to reduce the number of incorrect answers in favour of a 'no answer,' as a 'no answer' is considered to be better than a bad answer, we can find that Brill et al. incorporates a sophisticated QA pipeline in the AskMSR system in order to assess all the answers before display. The researchers are trying to go one step further by trying to predict which questions are not likely to return a correct answer. Thus, instead of losing time and unnecessarily increasing network traffic, the system would simply return a 'no answer' to such questions.

One study that is most relevant to our objective of creating a question answer system for the Mauritian Judiciary was conducted by Quaresma and Rodrigues [45]. As a matter of fact, our research shows that closed domain question answering system for the Legal domain is less popular. However, the research conducted by Quaresma and Rodrigues is of relevance to the aims of this project. Their research demonstrated that they aimed to have a question answering system for the legal domain with a view to widen and ease access to legal information. Evaluation of the system was done principally based on

180,000 Portuguese documents that consisted of Court decisions. The logic Programming, ISCO, was a very important technology for their system. Also, as numerous answers could exist for the same question, their system had the possibility of calculating the different answers that could exist and thereafter, they were returned to the system's user. The system was tested against 200 questions that were elaborated and tested by the same jury and results showed that the system could generate 50 correct answers, 22 imprecise answers and 125 bad answers. Joining the position of other authors like Matthew and Nyberg [46], Sasikumar and Sindhu [47] and Quaresma and Rodrigues also shared the view that for a question answering system to work effectively, especially in the case of a legal QA system, it is of utmost importance for the system to be rich in knowledge, have enough synonyms and to have a good representation of the domain. Good ontology is also considered as an important factor.

Another piece of work, which is of high relevance to our proposed objective, is that of Tirpude and Alvi [48]. The question answering system that has been proposed in this study is mainly a closed domain QA system that had as objective to retrieve legal documents pertaining to Indian Laws. The knowledge base used by the system was principally from legal documents of the Indian Courts. Their studies showed that keyword and metadata played a crucial part in retrieving information correctly. Through the testing and evaluation of their system, they also showed that higher accuracy is achieved where queries are structured as compared to a situation where a user of the system would ask random and unstructured queries that are in a non-question form. However, when testing their system against 100 random questions, the system could generate 62 correct answers. For achieving more precise answer generation, world knowledge was considered as more important than language knowledge and this was reaffirmed by Dillon [49] and Molla and Vicedo [50][51].

V. An Overview of the Different Existing Question Answer Systems and Information Retrieval Systems

The objective of this part is to give an apercu of the various question answering systems and information retrieval systems that exist. This part also elicits the main underlying differences and approaches of these systems. Given the lesser popularity of QA systems in the legal domain as enunciated by Molla and Vicedo [50][51], we shall not restrict ourselves to QA systems in the legal domain, but we shall also touch upon, both some of the other general purpose question answering systems as well as some other QA systems in the general domains. However, irrespective of whether it is a general purpose question answering system or a restricted one such as one in the legal, medical, architecture or life sciences, all these systems use different ways to process the input provided by users and to output the answers to the respective inputs. For the purpose of portraying the different QA and information retrieval systems that exist in both the legal and other domains and which are popularly used by legal practitioners, legal professionals, students and the Mauritian public at large, we propose to give an apercu of how the following systems work: the Supreme Court's of Mauritius website; Lexis Nexis; Google and Google Scholar. However, with respect to the general-purpose QA systems, START, AnswerBus, AskJeeves and Zapmeta shall be analysed. On the other hand, we shall also compare the general architecture of QA systems that exist in the medical field like AskHermes and MedQA.

A. The Website of the Supreme Court of Mauritius

The website of the Supreme Court of Mauritius serves as an important information system in the legal field especially for Judges, lawyers, attorneys, legal research assistants, students and the public at large. The website is commonly, if not daily, used by lawyers and attorneys for their legal research and to prepare their cases' arguments by referring to the legal precedents. The website of the Supreme Court is particularly limited to enable a user to find the various lists of judgments that have been pronounced by the different Courts of Mauritius namely: The Supreme Court, District Courts, the

Intermediate Court, the Industrial Court, the Court of Rodrigues and the Privy Council. The figure below illustrates the User Interface for the website of the Mauritian Supreme Court. The user will principally be able to choose the Court s/he wishes to obtain a judgment and the specific year in which the judgment was pronounced.

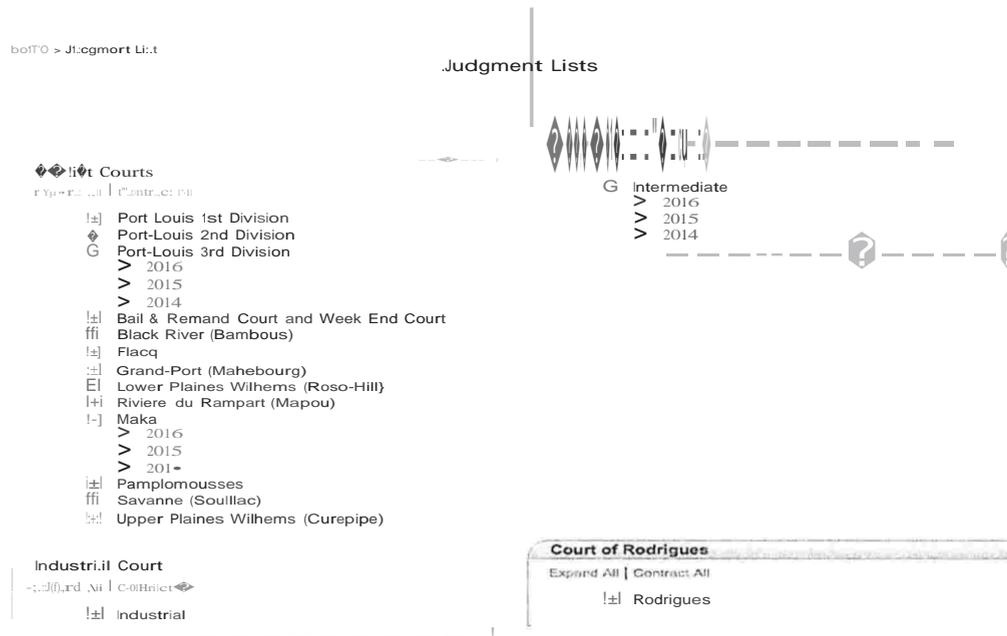


Figure 2

Source: <http://supremecourt.govmu.org/scourt/pubjudgment/pubViewP.do>

It can be seen from the above user interface that the options for retrieving information pertaining to a case is rather limited as there is at this point in time no such option to enable the user to find a document per area of law, example, cases related to breach of tenancy agreement or breach of contract. Unless the user know the name of the parties in the case, document number or date on which the judgment was delivered, it is quite cumbersome for the user to have relevant information about the case as shown in the snapshot below.

As shown by the above user interface, the document retrieval process is more fine-tuned. The user can here at the 'search by' option use the drop down menu and search a document by inputting a keyword, a subject, a summary or a title. By selecting 'by subject' for example, as shown in the picture below, the user is enabled to choose among a list of values in order to further refine the search.

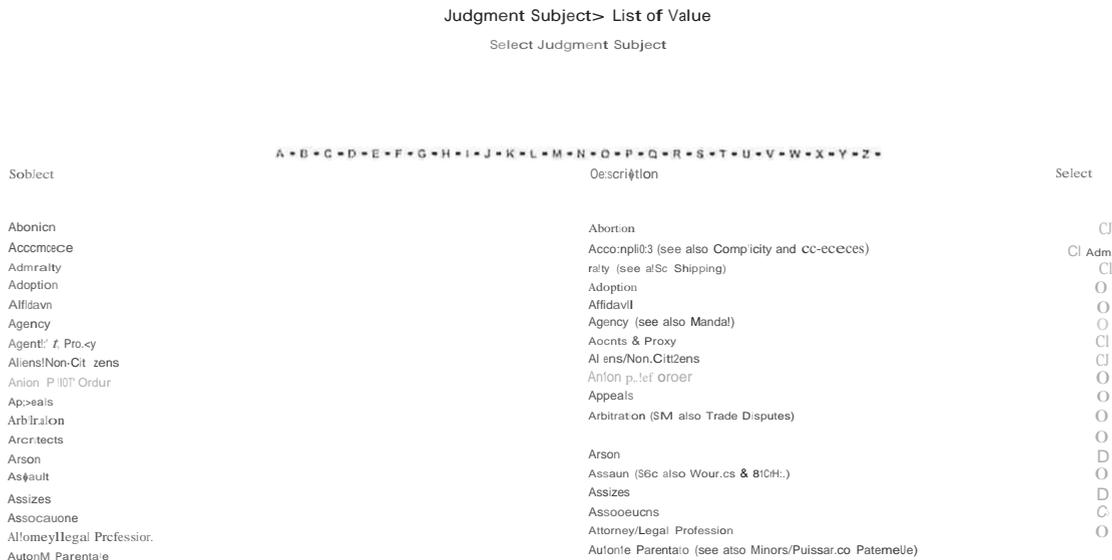


Figure 5

Source: <https://supremecourt.govmu.orn/scourt/judgment/searchAdvancedListP.do>

Arbitration, appeals, assault, banking law, company law, contract, damages, dangerous drugs are some of the list of values from which the user can choose. The list of values chosen is then submitted and thereafter, the user can from the interface choose the Court and the range of years for which s/he wishes to get a case which may be related to contract for example. A screen shot is shown below:

Search for document

Search By:

1. Select how to search from the drop down menu.
 2. You can include words in quotes (E.g. "contract") or use AND to clarify search string.
 3. You can search for a word.
 4. You can search for a word by using \ (backslash) before the 'enclosed between double quotes (E.g. "contract") as it is not enclosed between double quotes.
 5. You can search for a word by using * (asterisk) before the word to search for all words starting with that word.
 6. You can search for a word by using ? (question mark) after the word to search for all words ending with that word.
 7. You can search for a word by using # (hash) before the word to search for all words starting with that word.
 8. You can search for a word by using \$ (dollar sign) after the word to search for all words ending with that word.
 9. You can search for a word by using ^ (caret) before the word to search for all words starting with that word.
 10. You can search for a word by using _ (underscore) after the word to search for all words ending with that word.

Category: All ...

Period: From year: 2014 To year: 2016

1. Specify a search year or range of years. The default is the current year.
 2. You can specify a range of years by using the search filter in the response text boxes.

Your search returns 5 document(s)

Title	Summary	Document Number	Entered On
ESPACE MARIN V NUNKOO 17 F. Supp. 2d 1000, 1001 (D.C. Md. 2014) (Contract)		2014 INT 84	07/04/2014
KENDRA V BELGIAN QUARTER 17 F. Supp. 2d 1000, 1001 (D.C. Md. 2014) (Contract)		2014 INT 54	25/02/2014
FASL V MOOKLALL BEEGUN 17 F. Supp. 2d 1000, 1001 (D.C. Md. 2014) (Contract)		2014 INT 41	10/02/2014
D Contracting v Master Tech 17 F. Supp. 2d 1000, 1001 (D.C. Md. 2014) (Contract)		2014 INT 20	27/01/2014
IMC v Fide 17 F. Supp. 2d 1000, 1001 (D.C. Md. 2014) (Contract)		2014 INT 2	10/01/2014

Figure 6

Source: <https://supremecourt.govmu.org/scourt/judgment/searchAdvancedList.do>

From the above, we can see that the Supreme Court website provides for some interesting tools to search and refine search for a particular document retrieval as compared to a few years back, where such options were inexistent. However, it can be deduced that a whole list of documents is still returned to the user. The website is not yet capable for instance of generating precise answers or specific documents with respect to a search made.

B. Lexis Nexis

Lexis Nexis is also one of the most commonly used commercial tool in the legal profession and by the students' community for legal information retrieval. Access to Lexis Nexis is only possible after logging in and a subscription fee should also be usually paid by the user in order to benefit from its use. A first look at Lexis Nexis is quite impressive in terms of the different options it offers to the user. A look at its basic user interface shows that Lexis Nexis has the ability to extract information pertaining to cases, legislations, e-journals, forms and precedents, among others. This is shown below:

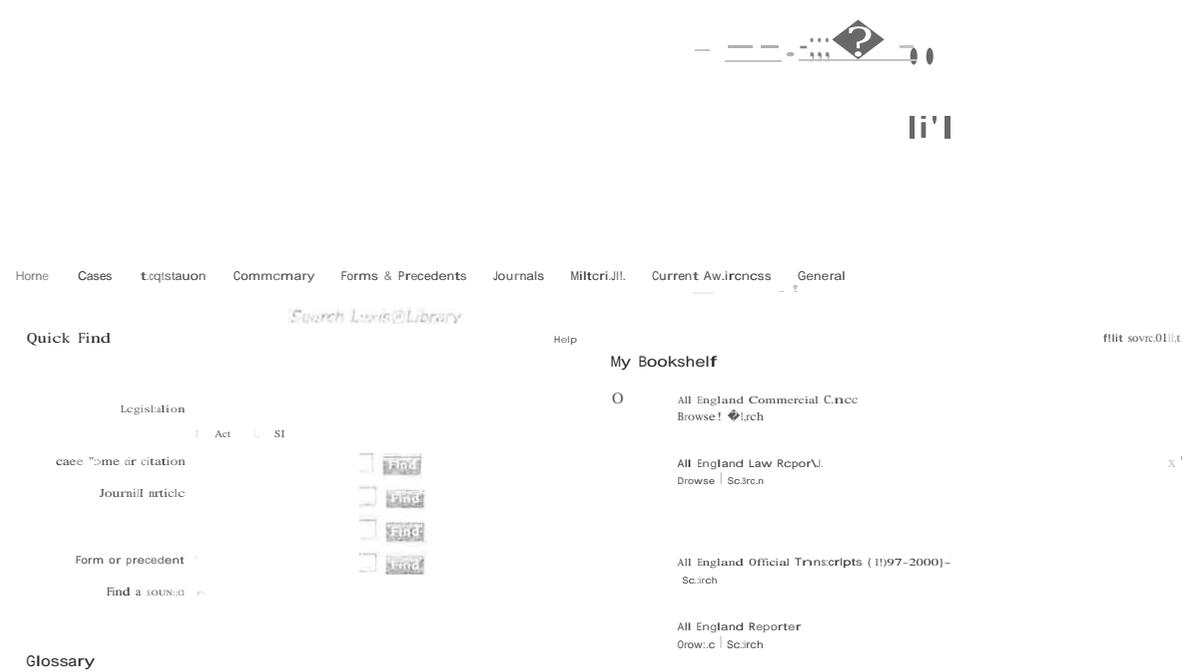


Figure 7

Source: <https://www.lexisnexis.com/uk/legal/auth/checkbrowser.do?jsessionid=D0297774AF0723BI836O47418EE9A355.iSgDUmefTxTFriW259uBg?t=1455350446573&bhcp=1>

However, for the purpose of this sub-part, we shall limit us to the different options that Lexis Nexis offers to the user for case' extraction and finding information pertaining to legislations. By clicking on the 'cases' icon, the user is presented with an interface where s/he is given the possibility of doing a basic search and an advanced search, which is optional as shown by the screenshot below:

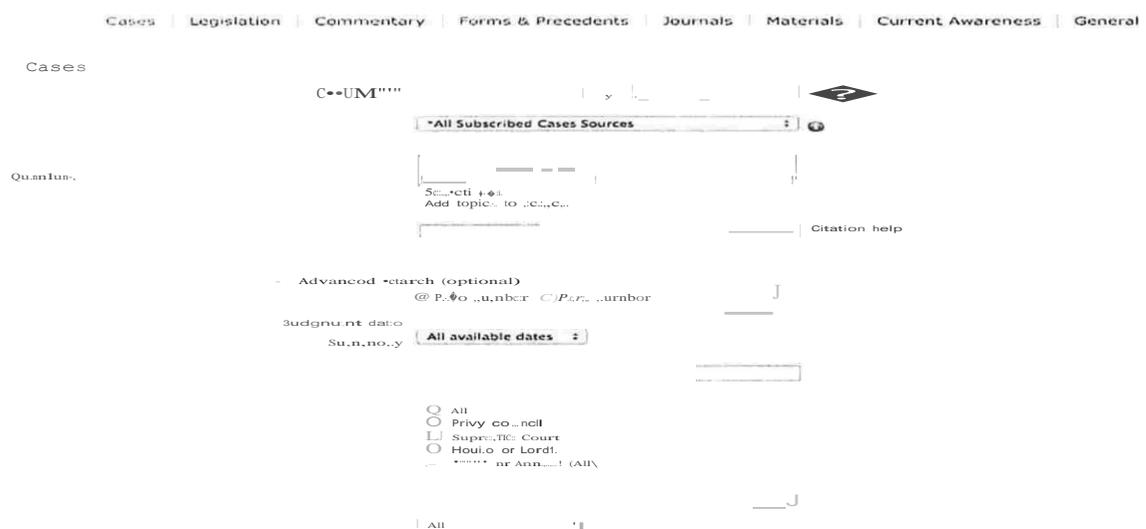


Figure 8

Source: <https://www.lexisnexis.com/uk/legal/>

J
J
U

Under the basic search, the user is able to look for a particular case by the parties' name and by its sources which may vary from 'All England commercial cases', 'All England Law Reports, All England Law Reports European cases, among others. Moreover and interestingly, under the 'search terms' option, the user is able to enter some key words that could further help Lexis Nexis to refine his/her search. For example, a user might not always have information about a particular case law. In such an instance, Lexis Nexis enables the user to locate cases that could be categorized under one particular area of law. For example, by entering "breach of tenancy agreement, in the white box of 'search terms,' the user is able to find cases that are related to 'breach of tenancy agreement' or where these words could potentially appear as shown in the figure below:

Lexis Recommends Hide ..

Key cases relating to Agreement • Identified by specialist law reporters and the All England team

Bayer AG v European Commission: C-2/01, C-1/01, C-1/01, C-1/01, [2004] All ER (EC) 500, 78 BMLR 91, [2004] All ER (D) 12 (Jan)
 Bayer AG v European Commission: T-1/96, [2001] All ER (EC) 1

Narrow By Sort by: [icon]

** Search within results

1 - 37 of 37 results shown

Source Name		Court/Tribunal	Judgment date	Source	Signal
All England Law Reports (7) Civil Overview (6) Estate Gazette Law Reports (5) All England Reporter (..)	1	Fareham Borough Council v Miller	CA, CIVIL DIVISION	6 March 2013	
<small>Reverts of flat being occupied by associates r/l defendant in breach of tenancy agreement • Incidents of anti-social behaviour at nat Defendant denying responsibility - 1 hit in the document</small>					
All England and Commonwealth Reports (1997-2005) (3) Show all 14 source Names	2	Fareham Borough Council v Miller - [2013] All ER (D) 53 (Mar)	Court of Appeal, Civil Division	6 March 2013	All England Reporter
<small>Reverts of flat being occupied by associates r/l</small>					

Figure 9

Source: [https://www.lexisnexis.com/uk/legal/search/casessubmitForm.do#0||JUDGMENT-DATE.D,J-L.\\$PSEUDOLOSK.A.HIII](https://www.lexisnexis.com/uk/legal/search/casessubmitForm.do#0||JUDGMENT-DATE.D,J-L.$PSEUDOLOSK.A.HIII)

However, under the 'Advanced search' interface, the user is given additional options which could further fine-tune his/her search for a particular case law. A user may under the '*judgment date*' dropdown choose when the case was decided as shown below:

The screenshot shows the Lexis Nexis search interface. On the left is a sidebar with navigation links: 'Cases', 'Case Overview', 'Quantum', and 'International Cases'. The main search area contains several input fields and options:

- Case name:** A text input field with a search icon.
- Sources:** A dropdown menu showing 'All Available Sources' and 'More sources'.
- Search terms:** A text input field.
- Citation:** A text input field with a 'Citation help' link.
- Advanced search (optional):** A section with radio buttons for 'Page number' and 'Para number'.
- Judgment date:** A dropdown menu with options: 'All available dates', 'Today', 'Date is before...', 'Date is after...', 'Date is between...', 'Previous week', 'Previous month', 'Previous 3 months', 'Previous 6 months', 'Previous year', and 'Previous 2 years'.
- Court:** A dropdown menu with options: 'Date is before...', 'Date is after...', 'Date is between...', 'Previous week', 'Previous month', 'Previous 3 months', 'Previous 6 months', 'Previous year', and 'Previous 2 years'.
- Judge(s):** A text input field.
- Reset list:** A button.

Figure 10

Source: <https://www.lexisnexis.com/uk/legal/search/flap.do?flap!D=cases&random=0.29596412990662546>

Moreover, s/he can also choose the Court dropdown menu where s/he could choose one or more Courts before which the case appeared. The cases can also be looked up by the name of the Judge(s) or Counsel who appeared in that case. With respect to finding a particular legislation, Lexis Nexis also offers the user some interesting tools that help him/her obtain the desired results. The user's interface for finding a legislation shows that users are able to select between the versions of the law. Very often, Parliament brings amendments to existing laws, make directives or regulations under an existing law. It is very important for those working in the legal domain, especially Judges, lawyers and attorneys to keep track of these changes especially when they have to prepare their cases and their arguments. This is a big barrier that legal professionals often face. There are at many times lack of information about all those historic developments, thereby making it difficult for them to be able to trace back any, if not all amendments which have been brought to a particular law. However, we can find that this option exists in Lexis Nexis. Indeed, the user may choose between getting the current version of the law or the historical version. The user may search by title, the year in which the law was enacted and may interestingly also look for a specific provision in a law instead of whole act.

C. Google v/s Google Scholar

Nowadays, Google is also used as the most common question answering system in the world. A difference is here made between the traditional Google interface and on the other hand, that of Google Scholar. Behind the white box appearing on Google's interface, there are several important steps behind the search query's ride. Indexing the web's content is the first crucial step in Google QA system. Ranking comes next and at this stage of the query, the system finds the most relevant answers to the user's query. Google also offers the option of auto-completion, where the user is given a forecast of what he might actually be typing in the white box. With respect to understanding the user's query, Google QA system will return the question asked to Google through the Internet. This is where indexing plays a major role. By parsing the question, Google will figure out what was the user's intent and using indexing, the system generates the best results to the user's query. At this stage, 'snippets' are also used to present to the user the most appropriate answers to the question asked.

On the other hand, Google Scholar is also a question answering system which answers to questions written in natural language. Google Scholar follows the same steps as Google for retrieving information. However, the Google Scholar interface is somewhat different to that of Google. It offers to the user some additional tools like 'Case Law' and 'Articles' and it also enables the user to do an advanced search where the results could be restricted to authors, subject area, publication date, among others. In terms of information presentation, Google Scholar shows the results that are mostly cited. However, with respect to determining citations for a specific article, Google Scholar is criticized for extracting a reference list that is short and it is also criticized as being capable of citing articles that are only electronically available [51] [52]. Despite being a great tool for legal professionals and students or for the public in general to carry out preliminary literature work or to obtain past cases, yet, Google Scholar is criticized as being limited in terms of editorial work and it does not contain statutes or regulations [53].

D. General Purpose Question Answering System (Open-Domain)

1. Zapmeta

This QA system is basically a meta-search engine which makes use of various search engines under one interface. It thus allows users to ask questions in a natural language and like Google, Zapmeta finds and ranks the most relevant answers to the user's query and it also offers the option of auto-completion where the user is given a forecast of what he might actually be typing as a query. This is shown below:



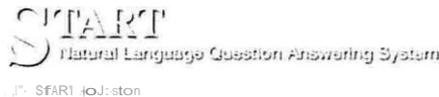
Figure 11

Source: <http://www.zapmeta.com>

Given that Zapmeta extracts information from several search engines, it avoids the user from having to spend time and search through the various search engines on an individual basis to find the answer to a particular query [54]. Thus, by having a meta-search engine, Zapmeta would give its results by using data from various existing engines.

2. START Question Answering System

START is also a question answering system which answers to questions in a natural language and which instead of returning a whole paragraph or text, it specifically answers to the questions asked. START is one of the very first QA system that is web based and which is successfully operating since December 1993 [55]. START comprises of a knowledge base against which it compares the queries obtained after breaking down the input by a user. In simple words, the software parses the query, a matching exercise is thereafter done where the question parsing is matched against the system's knowledge base. The relevant part of the query segment is then presented to the system's user [56]. START also uses an approach known as natural language annotation to unite the users to the information require [57]. By using this method, START is also able to deal with any kind of input ranging from text, pictures or even videos.



START, the world's first Web-based question answering system, has been on-line and continuously operating since December, 1993. It has been developed by Boris Katz and his associates of the InfoLab Group at the MIT Computer Science and Artificial Intelligence Laboratory. Unlike information retrieval systems (e.g., search engines), START aims to supply users with "just the right information," instead of merely providing a list of hits. Currently, the system can answer millions of English questions about places (e.g., cities, countries, lakes, coordinates, weather, maps, demographics, political and economic systems), movies (e.g. titles, actors, directors), people (e.g., birth dates, biographies), dictionary definitions, and much, much more. Below is a list of some of the things START knows about, with example questions. You can type your question above or select from the following examples, less...

Geography

- What South-American country has the largest population?
- What is the largest city in Florida?
- Give me the states that border Colorado.
- What cities are within 250 miles of the capital of Italy?
- How many people live in Israel?
- Show me a map of Denmark.

Science and Reference

- What is Jupiter's atmosphere made of?
- Who first discovered radiocarbon dating?
- How far is Neptune from the sun?
- Why is the sky blue?
- What planet has the smallest surface area?
- How many feet are there in a kilometer?

Figure 12

Source: <http://start.csail.mit.edu/index.php>

3. AnswerBus Question Answering System

The AnswerBus system is another open-domain QA system. It provides the user with the possibility of not only asking the question in English but in five additional languages as well. However, the output is given only in English. The system uses five main portals (search engines) in the quest to find the most suitable answers to the queries asked by a user. By using those portals, AnswerBus thereafter locates potential answers and sentences from the webpages. A description of how AnswerBus works can be grouped in five main steps. At the first step, the system accepts a question in natural language and verifies if it is in English. If not in English but is rather in one of the five other languages accepted by AnswerBus, the system uses a tool known as 'Babelfish' so as to translate the question in English. Three search engines are then chosen to obtain information for the query and then, the system models queries for the user input. AnswerBus then connects to the different portals to fetch the relevant information. At the fourth step, AnswerBus selects relevant sentences to form the answer to the question as formulated by the user. In the first place, the system, through parsing, breaks down the documents into sentences and through word matching, it assesses if a sentence is a potential answer to the user's query. At the last step, AnswerBus displays the sentences along with their links according to the order in which they are relevant [58].

4. Ask.com (Ask Jeeves)

Ask.com system, originally known as AskJeeves, possesses several portals. Like AnswerBus and START, this QA system allows the user to ask queries in natural language. After submission of a query by a user, Ask.com reformulates the query and matches the sentences against the web [59]. Answers are commonly answered on the fly or are extracted from a list of frequently asked questions. At the backbone, Ask.com basically works in the similar way as the other QA systems discussed above. One point worth noting however is that as compared to Google, which ranks the possible answer according to their popularity, Ask.com on the other hand, ranks answers according to topic communities and editorial functions [60].

E. Closed-Domain Question Answering Systems for the Medical Domain

Like it is for the legal domain where the amount of information and existing documentation is so large, similarly, the medical domain also possesses a very large amount of documentation and very often, professionals in the medical field make use of QA systems in order to obtain specific answers to their medical queries [61]. For the purpose of this sub-part, we analyse the approach and architecture of two common medical QA system: AskHermes and MedQA.

1. AskHermes

This QA system has the capacity of answering to complicated medical questions and to extract answers that are specific to the query asked [62]. Like the other QA systems discussed above, AskHermes enables the user to input the question in natural language. AskHermes' knowledge base is principally built from Wikipedia, documents from eMedicine, among others. What is however particularly singular to AskHermes is that the system categorises the queries in twelve general topics [61]. The aim of it is to make it easier to retrieve information. Moreover, with respect to the answer presentation, AskHermes uses mainly a clustering technique as questions asked in the medical field are often found to be long and related to several topics [63]. Instead of returning a list of documents based on ranking as output, as compare to other QA systems, AskHermes returns answer passages that are structurally clustered based on content-bearing query [62].

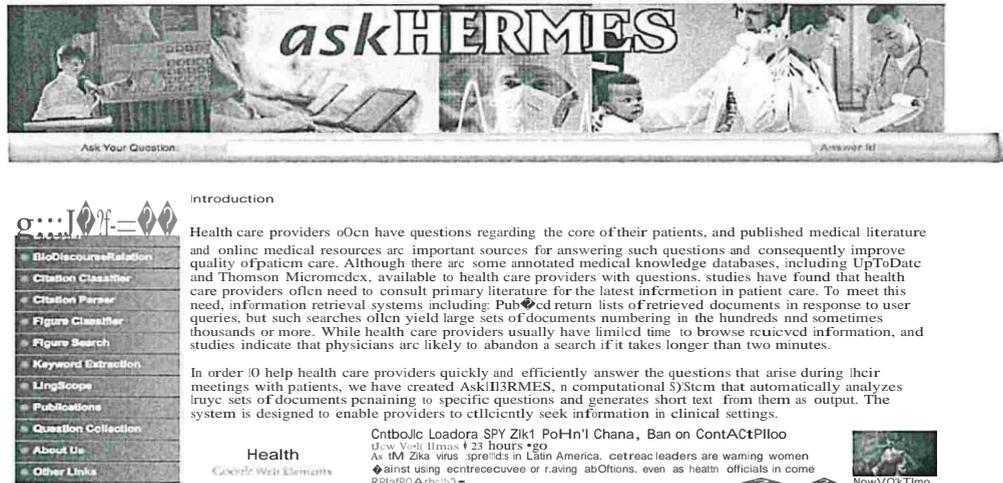


Figure 13

Source: <http://www.ask:hermes.org/index2.html>

2. MedQA

MedQA is also a popular QA system used by medical practitioners and is different from AskHermes with respect to question classification. The present system will classify the query asked by a user into a specific question type and for which a particular answer strategy will be used [64]. Moreover, documents are extracted principally from the web and from Medline. After extracting the answer, the system interestingly summarises the text obtained, whereby unnecessary sentences are dismissed. It then compresses the sentences into a logical summary. The user thus obtains a condensation of the sentences in the form of a summary.

F. Comparison of Answers by different Question Answering Systems

Table 2: Different ways in which different QA systems answer to complex questions

Sample Question: What is judicial review and what are the different stages in a judicial review process

Google	As answers we get several links that contain the answer principally to the first part of the question. However, with respect to the second part of the question, the user still has to open and go through the different proposed link to get information about the different stages in a judicial review process.
Google Scholar	In Google Scholar, the user, as compared to google.com, barely obtains information any information on the first part of the question. With the "Articles" option activated, Google Scholar also returns several links from different websites and through which the user has to skim through in order to obtain answer to the second part of the question. However, if the user clicks on the "case law" option on Google Scholar, s/he obtains a series of cases that relate to judicial review but which still do not provide precise answers to the question above.
Zap meta	Unlike both systems mentioned above, Zapmeta replied to both parts of the question in the very first answer it gives back and also returns several links which could contain the answer to both parts of the question.
START	Unlike Zapmeta, START replied perfectly to half of the question only while the second part remained totally unanswered. With respect to the question's first part, the answer is extracted from Wikipedia.
AnswerBus	As answers, we obtained several links, which could contain answers to the first part of the question but not to the second part.
Ask.com	Like AnswerBus, Ask.com also returned several links to the user, which could possibly answer to the question. By clicking on the first link returned by ask.com, we obtained a document that gives answers to both parts of the question asked but, the user still needs to open the links and go through the documents rather than obtaining a precise answer to the query.

By analyzing the output we obtained on the several QA systems; we can deduce that Zapmeta, Ask.com and Google.com are more reliable since the output we received matched our query to a certain extent. However, we can also say that though being reliable, the above systems are not perfect, especially in terms of accuracy of answers. One common aspect that could be found in all the above QA systems is that they all practically return several links which the user will have to open and search through in order to obtain a correct answer to his/her query. Ideally, the user would expect a QA system to have capacity of understanding and analyzing the question that a user has asked in natural language. However, after analyzing the outputs obtained from the different systems above, we can deduce that they are far from being able to find precise answers. Worst, at times, the systems return completely irrelevant answers.

The basic approach to answer questions in a question answering system is to apply natural language processing and information extraction. Currently, natural language processing techniques, POS tagging, WordNet, among others are being used in question answering systems [21]. Research has also found that the performance of the QA system relies on the number of possible solutions in the document. Upgrading question answering and answer processing are also regarded as two essential components in a QA system. Question processing comprises firstly of question classification and secondly, question reformulation. The former component points out the kind of question and its respective answer, while the latter component transforms the question in a format which is easily understood by the system. Answer processing however comprises of candidate answer ordering and filtering and a validation portion which facilitates communication with users.

VI. Implementation

In order to develop this question answer system, we had to learn how to use the following tools in an efficient and productive manner.

HTML5

HTML5 is the latest evolution of the standard that defines HTML with new elements, attributes, and behaviours, and a larger set of technologies that allows more diverse and powerful Web sites and applications. HTML5 has several advantages compared to the previous version. It provides better semantics which allowing us to describe more precisely what our content is.

CSS

Cascading Style Sheets, most of the time abbreviated as CSS, is a stylesheet language used to describe the presentation of a document written in HTML. CSS describes how the structured element must be rendered on screen.

PHP

PHP is a general-purpose programming language which is most often use for server-side programming. PHP is used to perform processing on the web server. PHP is one of the most important tools that have been in this project. PHP originally stood for Personal Home Page but now the meaning has changed to Hypertext Preprocessor.

MySQL

MySQL is a freely available open source Relational Database Management System (RDBMS) that uses Structured Query Language (SQL). SQL is the most popular language for adding, accessing and managing content in a database. It is most noted for its quick processing, proven reliability, ease and flexibility of use. MySQL is an essential part of almost every open source application. MySQL has been used as database server as it is free, is widely used and there are plenty of free tutorials to explain how to use it and what to do when errors are encountered.

The online Question Answer System has been developed using the above technologies.



Welcome to our Legal Support System.

To use this service, you should type in your queries or questions in the search box above and either press the Enter key or click on the Search button.

The system also accepts short phrases in double quotes, e.g., "consutution of mauritius", which may be followed by other keywords. Using double quotes may help you find the relevant information faster.

By default, two (2) results are displayed for queries. However, you may choose the number of results (1, 2, 5, 10, 20, 50) from the list provided.

If you wish to display only the names of acts which contain your keywords, you should tick the acts checkbox and then press the Enter key or click on the search button.

Currently, the system is limited only to few acts/statutes and few cases.

Kindly send us your feedback via email so that so this service can be improved for the benefit of legal practitioners, academics, students and even the public.

Thanks and regards
Project coordinator

Figure 14. Home Screen for the Question Answer System

Figure 14 above shows the main screen for the question answer system. One of the main objectives of our work is to provide a simple graphical user interface for users to interact with the system. The complexity will reside in the query processing and answer processing modules. The idea is to make the system useful not only to legal practitioners who know how acts, cases and other legal reports are structured and who has their own legal jargons to refer to legal entities but also to the laypersons who do not use legal terms. Our system allows a user to input his/her query using natural language, i.e., there are no specific rules to follow. He/She can type anything he/she wants in a free style.

Through a number of examples, we will explain how the system works and how to make the best use of it.

Example 1: Query in natural language

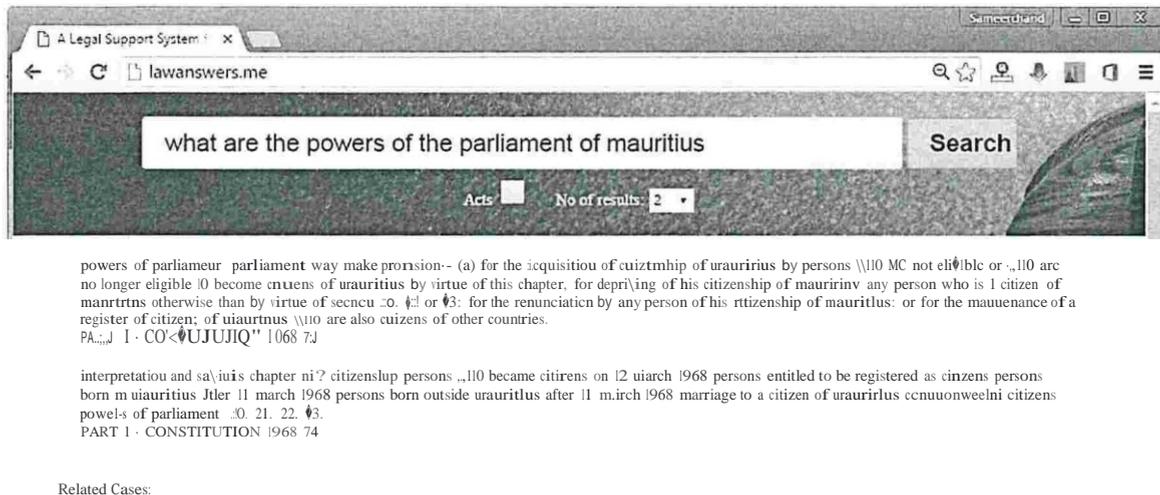


Figure 15. Sample Query 1 and Corresponding Answer

Figure 15 shows the answer returned by the system in response to the question *<what are the powers of the Parliament of Mauritius?>*. In this case, we have been very lucky in that the system has been able to return what can be considered as the best answer for this question. This answer comes from the Constitution of the Republic of Mauritius. The name of the act has been displayed below the answers. The user can click on the name in order to open it in a browser or download it to a computer or smartphone. Currently the system does not return the section number. Related cases are also returned. In this case, no answers have been returned. Our knowledge base consists of only few cases for the time being.

Our query processing module is able to extract the important terms from the query. For example, in this case, the keywords were *powers*, *Parliament* and *Mauritius*. The selected keywords are also highlighted in the answer. The system is case insensitive in that it makes no difference between lowercase and uppercase characters. The system returns two answers (by default) but the user can also choose to display any number of answers between one (1) and fifty (50) by choosing the relevant value from the *<No. of results>* drop down box..

Example 2: Query in natural language

The screenshot shows a search bar with the query "what are the duties of a company director" and a "Search" button. Below the search bar, there is a dropdown menu for "Acts" and a "No of results: 5" indicator. The search results are displayed as a list of text snippets, with the first three visible. The first snippet discusses the control of a liquidator by an official receiver. The second snippet discusses the duties of a director in certain cases. The third snippet, which is the best answer, is from Section 143 of the Companies Act 2001 and lists the duties of company directors in detail.

control of liquidator by official receiver where a person other than the official receiver is a liquidator or liquidator, the official receiver-- (a) shall take cognisance of his conduct and, if the liquidator does not faithfully perform his duties and duly observe all the requirements imposed on him, or contributory in that behalf, inquire into the matter and take such action as he thinks fit (b) if a complaint is made to the official receiver by a creditor or

may require the liquidator to answer any inquiry and provide any information or documents in relation to any winding up in which he is engaged: may apply to the court to examine him or any other person on oath concerning the winding up of the company; may direct an examination to be made of the books and vouchers of the liquidator: and may refer the matter to the director.

INSOLVENTY (ACT 1982)

sub-section (1) transactions involving self-interest meaning of "interested" disclosure of interest avoidance of transactions effect on third parties application of sections 149 and 150 in certain cases interested director may vote in certain cases of company information meaning of "relevant interest" relevant interests to be disregarded in certain cases disclosure of share dealing by directors restrictions on share dealing by directors sub-part 7? miscellaneous provisions relating to directors proceedings of board remuneration and other benefits standard of care and civil liability of officers indemnity and insurance duty of directors on insolvency sub-part 8? secretary registrar may approve firm of corporation for appointment as secretary qualifications of secretary duties of secretary notice to be given of removal or resignation of secretary part xii? enforcement interpretation of part xii sub-part a? injunctive 155.

COMPANIES ACT 2001

use of information and advice (1) subject to subsection (2), a director of a company shall, in relation to the matters concerned, rely on reports, statements, and financial data and other information prepared or supplied, and on professional or expert advice given, by any-- (a) (b) employee of the company whom the director believes on reasonable grounds to be reliable and competent in relation to the matters concerned: professional adviser or expert in relation to matters which the director believes on reasonable grounds to be within the person's professional or expert competence: other director or committee of directors upon which the director did not serve in relation to matters within the director's or committee's designated authority.

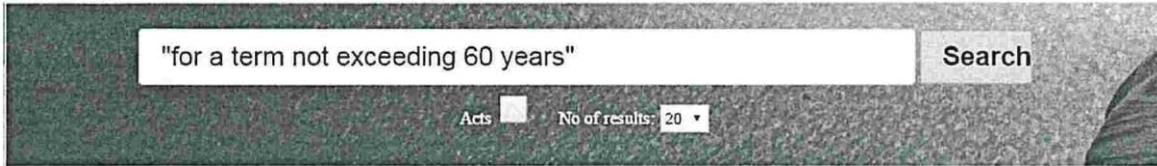
COMPANIES ACT 2001

Figure 16. Sample Query 2 and Corresponding Answer

Figure 16 shows the answer returned by the system in response to the question <what are the duties of a company director?>. Here also, the system has been able to extract the important keywords which are *duties*, *company* and *director*. These are highlighted in the answer. Five (5) results have been returned in this query, although only the first three have been displayed above. The third extract is the best answer as it comes from Section 143 of the Companies Act 2001 which list the duties of company directors in much details.

It is also good to point out that Section 143 is a very long section, as compared to others, and it is also one of the sections that are referred to the most often in relation to cases on the Companies Act 2001. Thus, it is important for our system to offer a good answer for such an important area of law. As future work, we can implement a facility to display the full section of the law to the user. The answers that are returned can also be ranked according to some criteria and then reordered, starting from the most relevant ones to the less relevant ones.

Example 4: Query as a keyphrase in double quotes



penalty for murder and infanticide (1) any person who is convicted of-- (>) murder or murder of" newly born child. shall be sentenced to penal servitude for life or, where the court is satisfied that substantial and compelling circumstances exist which justify the imposition of a lesser sentence and has entered those circumstances on the record of the proceedings, for a term not exceeding 60 years; attempt at murder or attempt at murder of a newly born child. shall be liable to penal servitude for life or, where the court is satisfied that substantial and compelling circumstances exist which justify the imposition of a lesser sentence and has entered those circumstances on the record of the proceedings, for a term not exceeding 60 years. CRIMINAL CODE 1833 7;

taking command of armed force any person, (a) (b) (c) without lawful right or reason, takes the command of an armed force, troop, fleet, squadron, man of war, fortified place, post, port or town: continues to hold any military command against the order of the government: or being a commander keeps his army or troops embodied after the disbanding or dismissing of the same has been ordered, shall be punished by penal servitude for life or, where the court is satisfied that substantial and compelling circumstances exist which justify the imposition of a lesser sentence and has entered those circumstances on the record of the proceedings, for a term not exceeding 60 years. CRIMINAL CODE 1833 71

plotting with foreign power (1) any person who does any plot or keeps up intelligence with any foreign power for the purpose of-- in order to move or stir it up to commit hostilities or levy war against the state, or to procure for it the means of so doing, shall be punished by-- (a) penal servitude for life or, where the court is satisfied that compelling reasons exist which justify the imposition of a lesser sentence and has specified the reasons on the record of the proceedings, for a term not exceeding 60 years; and forfeiture of property. CRIMINAL CODE 1833 7,

inducing or compelling president in the exercise of any lawful power any person who, by force or violence, induces or compels the president to exercise or refrain from exercising in any manner any of the powers vested in him by law, shall commit an offence and shall, on conviction, be liable to penal servitude for life or, where the court is satisfied that compelling reasons exist which justify the imposition of a lesser sentence and has specified the reasons on the record of the proceedings, for a term not exceeding 60 years.

CRIMINAL CODE 1838 71

62. stirring up civil war any attempt or plot, the object of which is to stir civil war by arming or by inciting the inhabitants to arm themselves against one another, or carry devastation, massacre or plunder into one or more districts, shall be punishable by penal servitude for life or, where the court is satisfied that compelling reasons exist which justify the imposition of a lesser sentence and has specified the reasons on the record of the proceedings, for a term not exceeding 60 years. CRIMINAL CODE 1833 73

inciting officer to mutiny any person who attempts to bribe, corrupt or seduce from his duty any officer or other person belonging to the state's land or sea forces, or attempts to incite him to mutiny or disobedience, shall be punished by penal servitude for life or, where the court is satisfied that substantial and compelling circumstances exist which justify the imposition of a lesser sentence and has entered those circumstances on the record of the proceedings, for a term not exceeding 60 years. CRIMINAL CODE 1838 73

arson causing death in all cases specified in section 346, where the fire caused the death of one or more persons, being in the place or vehicle set on fire at the moment of the fire breaking forth, the punishment shall be penal servitude for life or, where the court is satisfied that compelling reasons exist which justify the imposition of a lesser sentence and has specified the reasons on the record of the proceedings, for a term not exceeding 60 years. CRIMINAL CODE 1833 73

60. inciting citizen to rise up in arms every attempt or plot the object of which is to excite the citizens or other inhabitants of Mauritius to arm themselves against the state, shall be punished by-- (a) penal servitude for life or, where the court is satisfied that compelling reasons exist which justify the imposition of a lesser sentence and has specified the reasons on the record of the proceedings, for a term not exceeding 60 years; and forfeiture of property. CRIMINAL CODE 1833 71

stirring up war against the state any person who stirs up war against the state shall commit an offence and shall, on conviction, be sentenced to penal servitude for life or, where the court is satisfied that compelling reasons exist which justify the imposition of a lesser sentence and has specified the reasons on the record of the proceedings, for a term not exceeding 60 years. CRIMINAL CODE 1838 73

(3) notwithstanding subsection (2), any person convicted of an offence under section 29 (2) or 30 shall be sentenced to a fine not exceeding 2 million rupees together with penal servitude for a term not exceeding 60 years, where it is averred and proved that, having regard to all the circumstances of the case, the person was a drug trafficker. DRUGS ACT 1990 31

"master" has the same meaning as in the merchant shipping act: "maritime zones", "territorial sea", "internal waters", "archipelagic waters" and "historic waters" have the same meaning as in the maritime zones act: includes the sea: 3. piracy and maritime attack (1) (a) (b) acts 2011 1109 any person who commits an act of piracy or a maritime attack, shall commit an offence and shall, on conviction, be liable to penal servitude for a term not exceeding 60 years. PIRACY ACT 2011 14

(b) (c) the person being or exercising control of the ship under subsection (1), or doing the act under subsection (1), as the case may be, is a Mauritian citizen; the act is committed in Mauritius; or (5) any person who commits an offence under this section shall, on conviction, be liable to penal servitude for a term not exceeding 60 years.

the act is committed in Mauritius: Or acts 101 J 1153 (9) any person who commits an offence under this Section shall, on conviction, be liable to penal servitude for a term not exceeding 60 years.

Figure 18. Sample Query 4 and Corresponding Answer

An interesting feature of our system is its ability to extract exact phrases from the knowledge base. For example, suppose a user wanted to know the most severe sentences in the Republic of Mauritius. He/She can input the text *<for a term not exceeding 60 years>*. However, this requires that the user has some knowledge of the content and how it is structured. This feature will be very useful to those who will know how to use it well. In this specific case, thirteen (13) answers have been returned which contains the keyphrase, nine (9) of which are from the Criminal Code 1838, three (3) from the Piracy and Maritime Act 2011 and one (1) from the Dangerous Drugs Act 2000. Thus, at the click of a button, we are able to know which legislations contain the most severe sentences. In the same way, we can find all offences where the imprisonment is for a term not exceeding 30 years or any other keyphrases can be entered.

Example 5: Query as a keyphrase in double quotes with filter

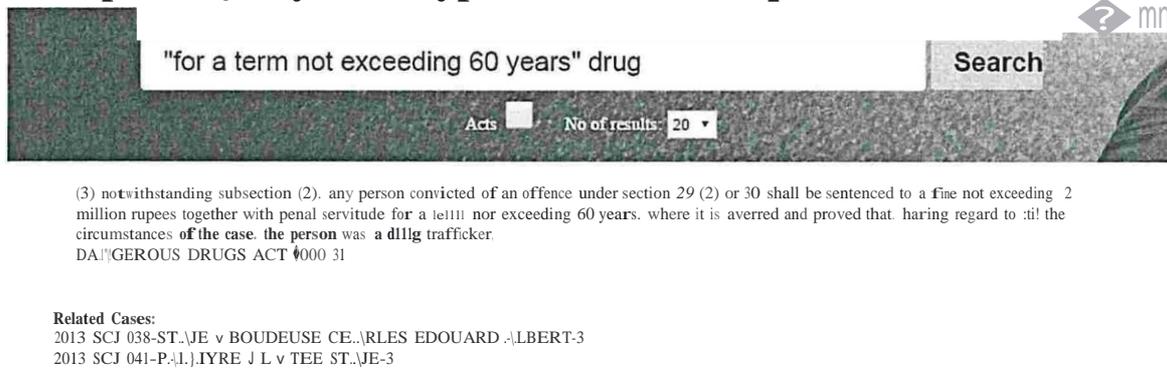


Figure 19. Sample Query 5 and Corresponding Answer

The system also offers the opportunity to refine a result if too many answers are returned. For example, in Example 4, thirteen (13) were returned. Now, if the user wants to get only those sections which deal with drug offences, he/she can enter the word *<drug>* as shown above after the double quoted keyphrase. Upon hitting search, only one answer is returned in this case. The relevant cases from the knowledge base are also returned.

Example 6: List of Acts



Am:
 L { TD IDUTIES AND TAXES ACT 1984
 LA., TILORD AC/D TEN...T ACT 1999 15
 MAURITIUS L.,\D TRANSPORT AUTHORITY ACT 1009 15
 E", TRO"> IDff A:, 1) L.,\CTI USE APPEAL TRJBU"1AL ACT 2010 05 "10T DI OPERATIO"
 LA'. 1) \, COUIS\TIOX \, CT 1973 II
 L.\ND SUR\TYQRS >CT 1076 00
 AG.\LEGA ISL.\, 1) COU"ICIL ACT 2004 01
 SHTE U"-TIS >C 1874 19
 OUTER ISL.\, 1)S DE\1:LOP\JE:ff CORPORATIO"1 ACT 198* 06
 STATE L.,\C1) \, \, LIEX.\TIOX\ ACT 1083 03

Related Cases:
 1013 SCJ 015-BUSJ).1:SS P.,RK OF \L.\URITIUS LTD., SOCJETE Fu"URA-4
 2013 SCJ 016-GR.\D PORT SA\,\<NI: DISTRICT COLJ"ICJL v \IONNERO"1\1 I\ &ORS-6
 2013 SCJ 016-DEERPAUL J v SEEBORUTH B-7
 2013 SCJ 030-TOOR.\B., LLY A ., D., BY A & ORS-6

Figure 20. Sample Query 6 and Corresponding Answer

In the example shown in Figure 19, we show that it is possible to list all the acts which contain one or more of the search keywords in their title by checking the **Acts** checkbox. For example, the keyword <land> returns 12 acts, ranging from Land (Duties and Taxes) Act 1984 to the State Land (Alienation) Act 1983. Four cases have also been returned. It is possible for the user to input more than one keyword or a keyphrase in double quotes.

Example 7: Display of Acts

LAND (DUTIES AND TAXES) ACT
 Act 46 of 1984— 16 July 1984

ARRANGEMENT OF SECTIONS

PART 1 – PRELIMINARY

1. Short title
2. Interpretation

Figure 21. Sample Query 7 and Corresponding Answer

A user can open the acts listed in any results in another window of the browser, as shown above in Figure 20, where the Land (Duties and Taxes) Act which was listed in Figure 19 has been opened for further reading. Once the act has been opened in a browser, the user can download the act, print it, perform additional search or share its link.

VII. Status of Outstanding Work from Progress Report Submitted on 08th March 2016

(a) Currently, the portal can only return only one answer in most cases. In the next version, it will be possible to display several answers for each query.

Update: It is now possible to display between 1 and 50 results for one query.

(b) The knowledge base itself has to be improved so that more relevant answers are retrieved and displayed to the user. For example, the section number should also be displayed.

Update: The results are fine in most cases but improvements can still be made. Displaying the section number for each answer requires some further work.

(c) A help section will also be included in order to explain to users how to use the system in the most efficient way.

Update: Guidelines have been provided on the home page on how to use the system.

(d) In the next version, we hope to make it possible for the user to download the full legislation from the answer.

Update: This is now possible.

(e) Currently the knowledge base consists only of few legislations only. We are planning to add some more legislations to the database.

Update: Some additional acts have been added but this is a continuous process.

(f) Judgements will also be added to the knowledge base in the next deliverable.

Update: Some cases have been added but this is a continuous process.

(g) Re-evaluate the portal after making the suggested improvements.

Update: Feedback was received from various stakeholders and the comments were taken into consideration as far as possible. For example, the user-interface is now more user-friendly and users can now access the portal using their smartphone.

(h) Write the final report.

Update: Completed.

VIII. Conclusions

This project is in line with national goals and the government vision of creating an eJudiciary system for the Republic of Mauritius. The portal eJudiciary.mu is such a platform which works 24/7 and which allows the submission of cases at anytime from anywhere. eJudiciary Mauritius is considered to be one of the most advanced electronic legal filing systems in the world and the first of its kind in Africa. It consists of components such as lodging of cases, filing of legal documents and other case management facilities. However, it does not cater for legal research. Thus, our web-based question answer and information retrieval system intends to bridge this gap. Legal research is a core activity in the legal domain and is both time-consuming and costly. The availability of information at the click of a button will assist the judiciary and other legal practitioners to become more efficient by hopefully help to reduce delays in the preparation of cases, writing of judgements and the disposal of cases. This system will contribute to making the human resources of the Mauritian Judiciary more efficient.

Publications

Pudaruth, S., Gunpath, R. P., Soyjaudah, K. M. S. and Domun, P., 2016. A Question Answer System for the Mauritian Judiciary. *Extended abstract submitted to the University of Mauritius Research Week 2016.*

Software

A web-based question answer system has been implemented.

The website can be accessed at: <http://lawanswers.me/>

Acknowledgements

This research project has been funded by the Mauritius Research Council.

REFERENCES

- [1] THE SUPREME COURT OF MAURITIUS, 2014. *Annual Report of the Judiciary*. Republic of Mauritius.
- [2] INNOCENT, T. and OBURA, C.O., 2010. An Assessment of Legal Information Needs and Access Problems of Lawyers in Uganda. *Library Philosophy and Practice [online]*, Paper 382.
- [3] OTIKE, J., 1997. Legal information needs of the general public with special reference to developing countries. *Library Review*, 46 (1).
- [4] LOGAN R.G., 1989. *Legal Information Sources in Law*. London: Gower.
- [5] MAYBURY, M. T., 2005. New Directions in Question Answering. *Journal of Computational Linguistics*, 31(3), pp. 413-418.
- [6] FAN, J., KALYANPUR A., GONDEK, D.C and FERRUCCI, D. A., 2012. Automatic Knowledge Extraction From Documents. *IBM J RES & DEV*, 56(3), pp. 1-10.
- [7] MCCARTHY, M.P and DENECKE, C.B., 2012. *Applied Natural Language Processing*. United States of America: IGI Global.
- [8] KWOK, C., ETZIONI, O. and WELD, D.S, 2013. Scaling question answering to the web. *ACM Transactions on Information Systems (TOIS)*, 4(4).
- [9] PEDRO, V., KO, J., NYBERG, E. and MITAMURA, T., 2004. An Information Repository Model for Advanced Question Answering Systems. In: *Proceedings of the 4th International Conference on Language Resources and Evaluation (CLRE)*, Lisbon, Portugal, 26-28 May.
- [10] Wikipedia, 2016. Question Answering [online]. Available from: http://en.wikipedia.org/wiki/Question_answering [Accessed 2 February 2016].
- [11] ATHIRA, P. M, SREEJA, M. and RAGHU RAJ P. C., 2013. Architecture of an Ontology-Based Domain – Specific Natural language Question Answering System. *international Journal of Web & Semantic Technology (IJWEST)*, 4(4).
- [12] COPPER, R. J. and RUGER, S. M., 2000. A Simple Question Answering System. In: *Proceedings of the 9th Text Retrieval Conference (TREC)*, Gaithersburg, Maryland, USA, 13-16 November.
- [13] TORRES-MORENO, J. M., ST-ONGE, P. L, GAGNON, M., EL-BEZE, M. and BELLOT, P., 2009. Automatic Summarization System coupled with a Question-Answering System (QAAS). Available from: <https://arxiv.org/pdf/0905.2990.pdf> [Accessed 2 February 2016]
- [14] KWOK, C., ETZIONI, O. and WELD, D. S., 2001. Scaling Question Answering to the Web. *ACM Transactions on Information Systems (TOIS)*, Vol. 19, No. 3.
- [15] PEDRO, V., KO, J., NYBERG, E. and MITAMURA, T., 2004. An Information Repository Model for Advanced Question Answering Systems. In: *Proceedings of the 4th International Conference on Language Resources and Evaluation (CLRE)*, Lisbon, Portugal, 26-28 May.
- [16] SHI, C., VERHAGEN, M. and PUSTEJOVSKY, J., 2014. A Conceptual Framework of Online Natural Language Processing Pipeline Application. In: *Proceedings of the Workshop on Open Infrastructures and Analysis Frameworks for HLT*, pp. 53-59, 23 August, Dublin, Ireland.
- [17] METAIS, E., 2002. Enhancing information systems management with natural language processing techniques. *International Journal of Computer Applications*, 53(4).

- [18] MUTHUKRISHAN, R. and SHANMUGASUNDRAM, H., 2012. A survey on question answering system. *International Journal of Research and Reviews in Information Sciences (IJRRIS)*, 12(1).
- [19] BILOTTI, M. W., 2004. *Query Expansion Techniques for Question Answering*. Thesis (PhD). Massachusetts Institute of Technology.
- [20] MONZ, C., 2003. Document Retrieval in the Context of Question Answering. In: *Proceedings of the 25th European Conference on IR research (ECIR)*, pp. 571-579, Pisa, Italy, 14-16 April.
- [21] WordNet, 2016. Wordnet.com [online]. Available from: <http://www.wordnet.com> [Accessed 2 February 2016].
- [22] KORENIUS, J., LAURIKKALA, J. and JUHOLA, M., 2004. Stemming and Lemmatization in the Clustering of Finish Text Documents. In: *Proceedings of the thirteenth ACM international conference on Information and knowledge management*, pp. 625-633, New York, USA.
- [23] HULL, D. 1996. Stemming algorithms: A case study for detailed evaluation. *Journal of the American Society for Information Science*, Volume 47(1), pp. 70-84.
- [24] TORRES-MORENO, J. M., 2012. Beyond Stemming and Lemmatization: Ultra-stemming to Improve Automatic Text Summarization. *CoRR Workshop on Service Oriented Computing*. Available from: <https://arxiv.org/pdf/1209.3126.pdf> [Accessed 2 February 2016].
- [25] GUPTA, P. and GUPTA, V. 2012. A survey of text question answering techniques. *International Journal of Computer Applications*. 53 (4).
- [26] SHARMA, S., PATANWALA, H., SHAH, M. and DEULKAN, K., 2015. A survey of medical question answering systems. *International Journal of Engineering and Technical Research (IJETR)*, 3(2).
- [27] JIBIN, F., KELIAN, J. and XU, J., 2009. Domain Ontology based automatic question answering. In: *Proceedings of the International Conference on Computer Engineering and Technology*, 22-24 Jan, Singapore.
- [28] BHOIR, V. and POTEY, M. A., 2014. Question answering: A heuristic approach. In: *Proceedings of the Fifth IEEE International Conference on Applications of Digital Information and Web Technologies*, 17-19 Feb, Bangalore, India.
- [29] WOODS, W.A., KAPLAN, R.M and WEBBER, B.N., 1972. *The Lunar Sciences Natural Language Information System*. Cambridge: Newman Inc.
- [30] TAHRI, A. and TIBERMACHINE, O., 2013. DBpedia based factoid question answering system. *International Journal of Web & Semantic Technology (IJWEST)*, 4(3).
- [31] ELENA, C., COJAN, J., APROSIO, A. P., MAGNINI, B., LAVELLI, A. AND FABIEN, G., 2012. QAKiS: an Open Domain QA System based on Relational Patterns. In: *Proceedings of the 11th International Semantic Web Conference (ISWC)*, 11-15 November, Boston, USA.
- [32] PEDOE, T., 2010. True knowledge: Open domain question answering using structure knowledge and inference. *AI Magazine* 31 (3), pp. 80-92.
- [33] BHOIR, V. and POTEY, M. A., 2014. Question answering: A heuristic approach. In: *Proceedings of the 5th IEEE International Conference on Applications of Digital Information and Web Technologies*, 17-19 Feb, Bangalore, India.

- [34] MOUSSA, A. M. and KADER R. A., 2011. QASYO: A question answering system for YAGO Ontology. *International Journal of Database Theory and Application*, 4(2).
- [35] SUCHANEK, F. M., KASNECI, G. and WEIKUM, G., 2008. YAGO: A large ontology from Wikipedia and WordNet. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 6(3), pp. 203-217.
- [36] Trueknowledge, 2016. True Knowledge. Available at: <http://www.trueknowledge.com/> [Accessed 3 February 2016].
- [37] DERICI, C., CELIK, K., KUTBAY, E., AYDIN, Y., GUNGOR, T., OZGUR, A. AND KARTAL, G., 2015. Question analysis for a closed domain question answering system. *Lecture Notes in Computer Science*, Vol. 9042, pp. 468-482.
- [38] FERRUCCI, D. A., 2012. Introduction to "This is Watson". *IBM Journal of Research and Development*, 56(3), pp. 235-249.
- [39] KOMIYA, K., ABE, Y., MORITA, H. AND KOTANI, Y., 2013. Question answering system using Q&A site corpus query expansion and candidate answer evaluation. *SpringerPlus*, Vol. 2, No. 396.
- [40] MOSCHITTI, A., 2003. Answer filtering via text categorization in question answering systems. In: *Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence*, pp. 241-248, 3-5 Nov, Sacramento, California.
- [41] PAKRAY, P., BHASKAR, P., BANERJEE, S. AND CHANDRA PAL, B., 2011. A hybrid question answering system based on information retrieval and answer validation. *CLEF Workshop on Question Answering For Machine Reading Evaluation (QA4MRE)*.
- [42] MENDES, A. C and COHEUR, L., 2011. An approach to answer selection in question answering based on semantic relations. In: *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, Vol. 3, pp. 1852-1857, Barcelona, Spain.
- [43] DONG, T., FURBACH, U., GLOCKNER, I. AND PELZER, B., 2011. A natural language question answering system as a participant in human Q&A portals. In: *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, Vol. 3, pp. 2430-2435, Barcelona, Spain.
- [44] BRILL, E., DUMAIS, E. AND BANKO, M., 2002. An analysis of the AskMSR question answering system. In: *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, Vol. 10, pp. 257-264, Stroudsburg, PA, USA.
- [45] QUARESMA, P. and RODRIGUES I. P., 2005. A question answer system for legal information retrieval. In: *Proceedings of the 18th Jurix Conference on Legal Knowledge and Information Systems*, pp. 91-100, 81-10 December, Brussels, Belgium.
- [46] MATTHEW, W. AND NYBERG, E., 2008. Improving text retrieval precision and answer accuracy in question answering systems. In: *Proceedings of the 2nd Workshop on Information Retrieval for Question Answering*, pp. 1-8, Stroudsburg, PA, USA.
- [47] SASIKUMAR, U. and SINDHU L., 2011. A survey of natural language question answering system. *International Journal of Database Theory and Application*, 4(2).
- [48] TIRPUDE, S. C and ALVI, A. S., 2015 Closed domain keyword based question answering system for legal documents of IPC sections and Indian Laws. *International Journal of Innovative Research in Computer and Communication Engineering*, 3(6).

- [49] DILLON, J., 1988. *Question and teaching: A manual of practice*. New York: Teachers College Press.
- [50] MOLLA, D. and VICEDO, J. L., 2007. Question answering in restricted domains: An overview. *Computational Linguistics Journal*, 33(1), pp. 41-61.
- [51] FALAGAS, E. M., PITSOUNI, E. I., MALIETZIS, G. A. AND PAPPAS, G., 2008. Comparison of PubMed, Scopus, Web of Science and Google Scholar: Strengths and weaknesses. *The FASEB Journal*, Vol. 22, pp. 338-342.
- [52] KOUSHA, K. and THELWALL, M., 2007. Google Scholar citations and Google Web/URL citations: A multi-discipline exploratory analysis. *Journal of the American Society for Information Science and Technology*, 58(7), 1055-1065.
- [53] The Bar Association of San Francisco, 2016. [online] Available at: https://www.sfbar.org/articles/google_scholar.aspx [Accessed on 7 February 2016].
- [54] MAMILLAPALLY, N. and MULUKULTA, T. 2013. Impact of meta search engines on information retrieval effectiveness: A case study. *International Journal of Advanced Trends in Computer Science and Engineering*, 2(6), pp. 118-123.
- [55] LIN, J. and KATZ, B., 2003. Question answering from the web using knowledge annotation and knowledge mining techniques. In: *Proceedings of the twelfth international conference on Information and knowledge management*, 3-8 November, New Orleans, Los Angeles, USA.
- [56] START, 2016. START Natural language question answering system [online]. Available from: <http://stmi.csail.mit.edu/start-system.php>. [Accessed on 8 February 2016].
- [57] MISHRA, A. and JAIN, S. K., 2016. A survey on question answering systems with classification. *Journal of King Saud University – Computer and Information Sciences*, 28(3), pp. 345-361.
- [58] ZHENG, Z., 2002. AnswerBus Question Answering System. In: *Proceedings of the 2nd International Conference on Human Language Technology Research*, pp. 399-404.
- [59] GUBANOV, M. AND PYAYT, A., 2012. MedReadFast: A structural information retrieval engine for big clinical text. In: *Proceedings of the 13th IEEE International Conference on Information Reuse and Integration (IR!*), 8-10 Aug, Las Vegas, USA.
- [60] MAZIN, G. AND FENG, J., 2008. *Speech and language processing over the world wide web*. *IEEE Signal Processing Magazine*, 25(3)
- [61] SHARMA, S., PATANWALA, H., SHAH, M. and DEULKAN, K., 2015. A survey of medical question answering systems. *International Journal of Engineering and Technical Research (IJETR)*, 3(2).
- [62] CAO, Y., LIU, F., SIMPSON, P., ANTIEUA, L., BENNETT, A., CIMINO, J. AND YU, H., 2011. AskHermes: An online question answering system for complex clinical questions. *Journal of Biomedical Informatics*, 44 (2), pp. 277-288.
- [63] BUI, Q. C. AND SLOOT, P. M. A., 2012. A robust approach to extract biomedical events from literature. *Bioinformatics*, 28(20), pp. 2654-2661.
- [64] HARIRI, N., 2013. Do natural language search engines really understand what users want? A comparative study of three natural language search engines and Google. *Online Information Review*, 37(2), pp. 287-303.

6
5
4

