2012

An ontology-based search framework to improve academic database retrieval for novice researchers

Norasykin Mohd Zaid
University of Wollongong

AN ONTOLOGY-BASED SEARCH FRAMEWORK TO IMPROVE ACADEMIC DATABASE RETRIEVAL FOR NOVICE RESEARCHERS

A thesis submitted in (partial) fulfilment of the
requirements for the award of the degree

DOCTOR OF PHILOSOPHY

from

UNIVERSITY OF WOLLONGONG

by

NORASYKIN MOHD ZAID
M.Ed. Educational Technology, Universiti Teknologi Malaysia
B.Sc. Computer (Multimedia) (Hons), Universiti Teknologi Malaysia

School of Information System and Technology
Faculty of Informatics
2012

THESIS CERTIFICATION

CERTIFICATION

I, Norasykin Mohd Zaid, declare that this thesis, submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Information System and Technology, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institution.

Norasykin Mohd Zaid
23 July 2012

LIST OF SYMBOLS AND ABBREVIATION

CD	Compact Disc
CLIR	Cross Language Information Retrieval
DBMS	Database Management System
EDR	Electronic Dictionary Research
ERIC	Education Resource Information Center
HTML	HyperText Markup Language
MARC	Machine-Readable Cataloging
MLIR	Multilingual Information Retrieval
MOHE	Ministry of Higher Education
MRD	Machine-Readable Dictionaries
NLP	Natural Language Processing
OPAC	Online Public Access Catalogue
OWL	Web Ontology Language
RDF	Resource Description Framework
SWQ	Semantic Web Query
TESL	Teaching English as a Second Language
TLA	Transactional Log Analysis
UKM	Universiti Kebangsaan Malaysia
UM	Universiti Malaya
UOW	University of Wollongong
URI	Uniform Resource Identifier
UTM	Universiti Teknologi Malaysia
W3C	World Wide Web Consortium
XML	Extensible Markup Language

LIST OF PUBLICATIONS

This is a list of conference papers and journal that are related to this research work:

Mohd Zaid, N. and Lau, S. K. (2008). 'Improving the Internet Search Capability by Semantic Technology', *in the 4th International Conference on Information Technology and Multimedia (ICIMU' 2008)*, UNITEN, Kuala Lumpur, Malaysia, 18th–19th November 2008, pp. 52-57.

Mohd Zaid, N. and Lau, S. K. (2011). 'Ontology Information Retrieval for Academic Resources', *in the International Business Information Management Conference (16th IBIMA)*, Kuala Lumpur, Malaysia, 29-30 June 2011, pp. 2273-2283.

Mohd Zaid, N. and Lau, S. K. (2011). 'Ontology Based Search Mechanism in Bilingual Database Resource', *in the Proceedings of the 11th International Conference Decision Sciences Institute and the 16th Annual Conference of Asia-Pacific Decision Sciences Institute (APDSI 2011)*, National Chengchi University, Taipei, Taiwan, 12th-16th July 2011, pp. 171-176.

Mohd Zaid, N. and Lau, S. K. (2011). 'Development of Ontology Information Retrieval System for Novice Researchers in Malaysia', *Journal of Software and Systems Development*, IBIMA Publishing, pp. 1-11.

ABSTRACT

This research investigates the development of an ontology-based search framework to assist inexperienced research students who are novice researchers in identifying research topics. The cohort of inexperienced research students face problems of limited research experience in conducting search queries, which often result in obtaining irrelevant search results. Other problems encountered by novice researchers include having difficulty in using the right keyword to search; not having any research ideas when conducting searches; irrelevant results returned; and the need to make multiple queries in a multi-language context. Advancements in the Semantic Web and ontology presents an opportunity to help novice researchers in overcoming the difficulties as described above. The database ontology can be mapped from the existing relational database to enable search queries to be conducted in a semantic context. This research proposes a semantic search approach that applies an ontology-based framework to improve the search process. The rationale behind the approach is based on the human perception of organising concepts in a proposed domain by transforming the structure of the concept into a mind-map. The elements of a mind-map are arranged according to major/broad concepts and can be narrowed down to the minor/specific elements of hierarchical structure. This approach represents a human thinking process in categorising categories-subcategories items in real life applications. Thus, a mind-map can be applied to construct search options in the search process. A system prototype has been developed to demonstrate the applicability of the proposed ontology-based approach. A survey has been conducted to evaluate the prototype and results have indicated that respondents found this approach to be useful in the search process. The contributions of the proposed approach are two-fold. Firstly users can formulate the query based on concepts and inter-relationships of concepts in the domain knowledge using a mind-map. Secondly the ontology-based approach improves the information retrieval process.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the guidance and help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, my utmost gratitude to Dr Sim Kim Lau whose sincerity and encouragement I will never forget. She has guided me with her patience and professionalism that greatly helps me and motivates me to continue and complete the thesis. Appreciation and gratitude also goes to all members of the School of Information System and Technology, School of Informatics, University of Wollongong for the great help and support.

Special thanks to my university, Universiti Teknologi Malaysia, and the Malaysian government (Ministry of Higher Education) for the financial sponsorships as well as the Faculty of Education and Department of Multimedia that was involved in the research evaluation process. I also wish to extend my thankfulness to all my colleagues and friends around me who have supported me throughout my PhD candidature with their love, patience and understanding.

I would also like to thank my mum and family members for always supporting me and encouraging me with their love and best wishes.

Finally, I would like to thank my husband, Mohd Fauzi, and my children for always being there and who stood by me through the good times and hard times.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The widespread proliferation of the Internet and the World Wide Web (WWW) in the last ten to fifteen years has resulted in using the WWW as an information-storing and information-searching medium. The WWW is a collection of human knowledge available in one hub to be shared by others in a wide range of topics and areas (Berners-Lee *et al.*, 1994). It has become an important information space of data repository and services. The main reason for people using the WWW is to find information that is of interest to individual needs and interests. One of the problems often encountered by information seekers is dissatisfaction in not being able to find desired information and users become dissatisfied with the results returned. Human judgment is often required in determining which searched documents are relevant and which are not. Therefore, merely using the WWW as a medium to search for information will not necessarily result in obtaining the desired information the users are looking for.

Advances in the Semantic Web and ontology provide an opportunity to overcome keyword search problems to allow meaning of words to be associated. For example Zaire is the former name of Democratic Republic of Congo. If users have no knowledge of the historical country names, their search for the Democratic Republic of Congo may not find documents about Zaire, resulting in incomplete information being returned. A number of valuable studies of web searches apply the ability of Semantic Web technology to overcome limitations of conventional database search systems (Ding *et al.*, 2004, Davies *et al.*, 2004, Finin *et al.*, 2005, Song *et al.*, 2005, Straccia and Troncy, 2006, Kohler *et al.*, 2006, Quan, 2007, Qasem *et al.*, 2008). The limitations are due to the inability of the system to search for documents using a static database structure. Some studies (Levy *et al.*, 1995, Yigal Arens *et al.*, 1996, Chiang *et al.*, 2001) have reported on the use of ontology-based searches to overcome the above limitations of conventional database search systems. Chiang *et al.*, (2001) summarises four approaches to search for unstructured web data: (i) syntactic search (Yuwono and Lee, 1996, Aggarwal *et al.*, 1998), (ii) metadata search (Lim and Ng, 1997, Martin *et al.*, 1998), (iii) query-by-example (Cascia *et al.*, 1998,

Swain, 1999), and (iv) navigational search (Menczer and Belew, 1998, Luah *et al.*, 1999). Syntactic search is defined as a non-structured or semi-structured data that can be experienced in the existing search engines. Metadata search is similar to the process used by syntactic data but with search enhancement on syntactic elements of documents such as titles, section headings and links. Incorporated in the metadata search is semantic metadata, which usually is performed in domain specific knowledge to enhance precision and recall of the query made. Query-by-example is used for non-textual searches such as images by expanding the example document tool to refine the search. In a navigational search, an agent is used by traversing through the links from sample web pages to find relevancy based on search specifications.

This research aims to investigate an ontology-based search approach to overcome limitations of search and retrieval problems in conventional database systems which include static structure, lack of query relevancy, limitations on query formulation and results relevancy. The remainder of this chapter is organised as follows: Section 1.1 presents a statement of problems. Section 1.2 states the research aims. Section 1.3 presents research objectives followed by section 1.4 which describes research methods. Section 1.5 outlines the research contributions and organisation of the thesis is presented in Section 1.6.

1.1 Statement of Problems

Searching for information can be difficult to perform when a complex search is conducted. Retrieval problems in information searching and retrieval can be categorised as: (1) query formulation, (2) lack of user perseverance, (3) terminology and search terms, (4) lack of user experience, and (5) synonyms and homonyms. These problems arise due to the need to describe the meaning of the search words. Semantic search is believed to have the potential to solve problems by being able to handle heterogeneous information resources, be adapted in the information domain to help making a search, filtering and ranking results efficiently, understanding the meaning of data and retrieve documents with complex relationships. Ontology provides a means in which semantic searches can be implemented. Ontology

databases can improve search result accuracy as the structured categorisation of classes and subclasses enables contextual relationships to be defined and data is given a well-defined meaning that is consistent across context. Thus, semantic searching that is capable of understanding the search intent can overcome failing of traditional search approaches.

This research proposes an ontology-based search framework to assist users in formulating search queries. The research uses a case study approach of developing an online ontology-based search system for a Malaysian university. The proposed system is developed as a resource to help students to search for past thesis reports for students at the Education Faculty of Universiti Teknologi Malaysia (UTM) which is one of the public universities in Malaysia. The current search system uses keyword searches to match the title of the thesis in the general category that includes Teaching and Learning, Science and Mathematics, Sports Science, Module and Education Studies. These categories have been determined by the librarian, based on major subjects offered in the faculty. Students can also browse for thesis titles using the supervisor's name and the year the thesis was completed. Common problems encountered by students when using this kind of search system include: (1) having difficulty in knowing the right keyword to use, (2) search results do not match with desired disciplines or topic areas (Mohd Zaid and Lau, 2008), and (3) search results often not matching the expected search outcome. When performing an online search, students usually have a preconception or an idea of a broad topic that they would like to find. However, they may not know how to formulate their idea into a search query to obtain the desired results. Thus, search results can sometimes return results which are not of their interests or meet their expectations.

Another problem that has been identified in the current database search system in the university is that a search cannot be conducted for a cross-disciplinary thesis. This is due to the structure of the Education Faculty which consists of five academic departments and each department has its own field of study and research areas. Currently, the theses are categorised based on academic departments only. However, theses can be supervised across discipline areas. For instance, a thesis that investigates computer-based teaching methodology in a chemistry subject can be

undertaken by a student enrolled in the Science department, who is studying for a teaching diploma course in the Education faculty, under the supervision of an academic in the Multimedia department. If the students were to search for a past thesis based on an academic department they may not find the multi-discipline thesis that is not categorised under the department that the student searches for.

An additional problem that has been identified in the current database search system is that a search can only be conducted in a mono-language (for example: Bahasa Malaysia or English only). Although most of the theses were written in the native language of Bahasa Malaysia, there are students who will write their thesis in the English language such as international students or students who are enrolled in Teaching English as a Second Language (TESL) course. As a result, students who conduct searches using keywords in the native language of Bahasa Malaysia may not find a thesis that is written in the English language and vice versa. In this case, the search needs to be performed twice, once in the native Bahasa Malaysia language and another using the English language if students want to ensure a complete search is conducted on all thesis records in both languages.

1.2 Research Aims

With the above problems identified, this research seeks to investigate the application of ontology-based search framework to overcome query formulation problems. The aim of this research is to propose and develop an ontology-based search framework that can assist novice researchers in formulating effective queries to enhance search processes. The research uses the case study approach to develop a prototype system to demonstrate the application of the proposed framework.

1.3 Research Objectives

The objectives of this research are as follows:

1. To propose an ontology-based search framework that assists in query formulation.
2. To analyse the effectiveness of the ontology-based search framework to help novice researchers in searching titles of theses.

1.4 Research Methods

This research is conducted in the following stages:

1. Literature review is conducted on methods and issues relating to information and database search and retrieval. Current limitations in online database search systems for novice researchers are presented to identify research gaps of current retrieval methods. Analysis of current limitations resulting in five groups of problems being categorised as: (1) query formulation, (2) lack of user perseverance, (3) terminology and search terms, (4) lack of user experience, and (5) synonyms and homonyms. Literature on Semantic Web and ontology are then conducted. Methods on constructing ontology are also presented by investigating different approaches of constructing ontology using ontology mapping and merging. Literature on the application of Semantic Web in information retrieval is also being analysed.

2. Development of a theoretical framework that aims to overcome problems faced by novice researchers will be conducted by proposing an ontology-based search framework to help in query formulation.

3. A design science approach is used to describe the application of the theoretical framework to a case study of ontology-based search systems for a public university in Malaysia. The design science approach is adopted based on Ken *et al.* (2007). There are six steps in this approach: (1) problem identification and

motivation, (2) objectives of solution, (3) design and development, (4) demonstration, (5) evaluation, and (6) communication.

1.5 Research Contributions

The original contribution of this research is the development of a theoretical framework of ontology-based search that can assist novice users in query formulation. The expected advantage of this approach is that users can easily formulate a query in which the query hierarchical structure is designed, based on how users will categorise the query keywords that are similar to mind-maps, that allows the query to begin with a broad category and narrowed down to a subcategory.

1.6 Organisation of Thesis

In this section, a brief description of each chapter is given.

Chapter 1 : An overview of thesis research work is presented through a statement of problems, research aims and objectives, research methods and research contributions.

Chapter 2 : In this chapter, a literature review is presented on background issues and problems of information search and retrieval. Major database search problems are identified. Cross language and multilingual information retrieval and background on the Online Public Access Catalogue (OPAC) system are presented. This is followed by literature on Semantic Web and ontology technologies. Approaches on ontology construction and Semantic Web search applications are also discussed.

Chapter 3 : Major information retrieval issues are presented in Section 1 of this chapter. The major issues identified are: (1) query formulation, (2) lack of user perseverance, (3) terminology and search terms, (4) lack of user experience, and (5) synonyms and homonyms. The chapter

discusses the development of an ontology-based search to assist novice users in query formulation. Application of the proposed framework is illustrated with an example.

Chapter 4 : This chapter presents the application of the ontology-based search framework to a case study. A prototype system is developed by using the design science method proposed by Ken *et al.* (2007) which consists of the following six steps: (1) problem identification and motivation, (2) objectives of solution, (3) design and development, (4) demonstration, (5) evaluation, and (6) communication.

Chapter 5 : This final chapter summarises research findings, research contributions, research limitations and future research direction.

CHAPTER 2: LITERATURE REVIEW

This chapter discusses the literature in the area of information and database search and retrieval. In this chapter, literature review is presented on background issues and problems of information search and retrieval. Major database search problems are identified. Cross language and multilingual information retrieval and background on the Online Public Access Catalogue (OPAC) system are also presented. This is followed by literature on Semantic Web and ontology technologies. Approaches on ontology construction and Semantic Web search applications are also discussed.

The chapter is organised as follows: Section 2.1 discusses information and database search and retrieval in online search systems. Search and retrieval approaches are discussed as well as search techniques and retrieval models. In addition, differences between searching and browsing are explained and the effectiveness of various information retrieval methods is discussed. Section 2.2 explores problems in information and database searches by categorising the issues into five categories: (1) query formulation, (2) lack of user perseverance, (3) terminology and search terms, (4) lack of user experience, and (5) synonyms and homonyms. Section 2.3 describes approaches in cross language and multilingual information retrieval. Section 2.4 presents general characteristic of the OPAC system with examples of an academic library in Malaysia being described. In Section 2.5, Semantic Web and ontology technologies are discussed. Section 2.6 focuses on ontology concepts including approaches of ontology construction and Semantic Web search applications. Finally, a conclusion is presented in Section 2.7.

2.1 Information and Database Search and Retrieval

The concept of information retrieval was introduced when we began to store information in computerised databases. An information retrieval technique is required to retrieve the required records when needs arise. Research has been conducted to improve capabilities and facilities in the area of information search and retrieval (Bates, 1989, Klink *et al.*, 2004, Malki *et al.*, 2008). The aim of the

information retrieval process is to help users to search and use information efficiently (Becker and Hayes, 1963). Large *et al.* (1999) describes the aim of information retrieval is to match information needs with items in the database. Extensive research has been conducted in the field of information retrieval to assist people in information seeking and various definitions of information retrieval have been defined (Becker and Hayes, 1963, Hancock-Beaulieu, 1990, Hancock-Beaulieu, 1993, Korfhage, 1997, Jansen and Pooch, 2001, Jansen and Spink, 2004, Marchionini *et al.*, 2004, Marchionini and White, 2007, Chowdhury *et al.*, 2010). Baeza-Yates and Ribeiro-Neto (1999) describe information retrieval as a process of how information is being stored, organised, represented and easily accessed by the user. While Chowdhury (2004) describes information retrieval as retrieving documents to meet the user's scope of search. Meadow *et al.* (2007) define information retrieval as a specific condition that concerns searching and retrieving of information stored in a database. However, Baeza-Yates and Ribeiro-Neto (1999) differentiated the terms between information and data retrievals, where information retrieval usually deals with unstructured information in its own original language, while data retrieval deals with well structured data such as relational database.

Kagolovsky and Moehr (2003) propose that information retrieval be defined by four characteristics: (1) consisting of structural and functional characteristics, (2) a dynamic iterative process, (3) assess by satisfying user's information needs, and (4) information retrieval taking place in environments that use the library databases. Therefore, information retrieval can be seen as a series of processes involving communication between users and systems to obtain desired information. Baeza-Yates and Ribeiro-Neto (1999) describe the process of search using the following six steps: (1) specify user needs, (2) formulate query operations, (3) process query, (4) obtain retrieved documents, (5) rank retrieved documents, and (6) analyse user feedback to alter query formulation (if required). Meadow *et al.* (2007) also lists the following steps of the information query process which includes: (1) interpretation of a query, (2) extraction of terms and indexes search, (3) identifying matched records, (4) confirming list of records matched, (5) displaying formatted output, and (6) user receiving the outputs. Marchionini and White (2007) explain the five steps in information retrieval as: (1) recognise, accept and formulate the problem, (2) express

the problem, (3) examine results, (4) reformulate the problem, and (5) use information.

Retrieval approach is another vital topic in the information retrieval area. The choice of approach depends on what kind of information the searchers have to start their search. Chu (2003) claims that when the searchers know exactly what they are looking for and the search query is well-defined, then searching becomes more efficient. Large *et al.* (1999) claim that when searching, searchers aim for a fair number of documents retrieved in search results, not seeing too many irrelevant documents in the search results, and to see as many relevant documents appear in the search results. Browsing is an alternative choice of seeking information when searchers find difficulties in identifying the exact information they need. Chu (2003) lists several situations where people like to browse (rather than search), which includes searching topics that are not clearly described; information cannot be stated clearly; to get a general idea of the information content; selecting information from a group of relevant and irrelevant items; and discovering new information. Large *et al.* (1999) classifies the ability of browsing to help find information which is ill-defined and to have only general information of topics searched for. Toms (2000) claims that the act of browsing relies on the available text that represents items to be assessed and connected by the searcher.

2.1.1 Information Retrieval Models

Various models of information retrieval have been developed over the years. These models are intended to facilitate an understanding of the retrieval process and have led to a formulation of different types of query techniques. The Boolean logic model is the first model developed by George Boole in the mid-19th century and has been commonly used in current search techniques (Chu, 2003, Chowdhury, 2004). This model supports three logical operations, which use the *AND* operator for the logical product (x), the *NOT* operator for the logical difference, (-) and the *OR* operator for the logical sum (+). Nevertheless, users have encountered many limitations while searching based on the Boolean model. Chowdhury (2004) identifies three limitations which includes formulation of search statements, number of retrieved items, and retrieved items cannot be ranked. Firstly, users may encounter problems

when attempting to put together a search query with the combination of logical operators such as *AND*, *OR*, and *NOT* which may result in outputs that are too narrow or too broad. Secondly, searchers do not know how many records are to be retrieved when a query is sent. In other words, searchers cannot predict if the search statement can achieve the right number of results. The third limitation is that search results cannot be ranked appropriately because the documents are identified as relevant by matching to the query term, thus ignoring the documents' rank of relevancy. Consequently, the vector space model was proposed to overcome some of the Boolean logic model limitations. This model can improve the retrieval performance by its term-weighting scheme, which can improve retrieval performances. Its partial matching strategy allows the matched documents to be retrieved based on a query statement, and its cosine ranking formula sorts the retrieved documents based on relevancy of similarity to the query terms (Baeza-Yates and Ribeiro-Neto, 1999).

Chu (2003) identifies three unsolved problems in the Boolean search that cannot be catered for by the vector space model, which are: (1) unable to define relationships for terms, (2) unsupported Boolean and proximity operators, and (3) the weighting process becomes subjective and complex. The first limitation is described as not being able to express meaning that exists between search terms. This problem cannot be solved even when vector space model is used to convey any kind of relationship that exists between search terms. While the second limitation is making it difficult for searchers to run the search as they cannot apply the *OR* operator or even the *WITH* operator, as this model does not support the Boolean and proximity operators. Apart from these two limitations, the third limitation is more difficult and the decision is subjective where searchers have to assign the terms' weights based on judgement.

Thus, a probabilistic model is proposed to determine the relevance between queries and documents it retrieved (Chu, 2003). This model is intended to estimate the probability that the required documents are relevant to the query term. However, Baeza-Yates and Ribeiro-Neto (1999) identify the following disadvantages of the probabilistic model which include: (1) no categorisation of documents whether they

are relevant or not relevant, (2) not considering document weighting in building the index term, and (3) freedom of supposition for the index term. In general, the idea of a probabilistic model is within a probabilistic scope, which allows the user to retrieve and categorise which documents are relevant and which documents are not. This process needs to be repeated until a certain stage where a set of answer patterns can be used to describe the query made. Unfortunately, the probability that those retrieved documents are relevant to the query cannot be computed, making it one of the weaknesses in the probabilistic model. Secondly, a weakness identified in the probabilistic model is how this model has set the relevancy of documents by only assuming the weights of relevancy with binary data, which are not compliant to the frequency of index terms, which appear in each document. As there is yet any documents retrieved in the first query attempt, index terms are often set to 0.5 as an initial guess value (Baeza-Yates and Ribeiro-Neto, 1999). Under the probabilistic model, as the document relevancy is important to determine the accuracy of documents relevant to the query, the probabilistic ranking principle is used to overcome this weakness, which regards document-by-document information retrieval in which results are assessed request-by-request, once the rank is being satisfied by the user (Robertson, 1977, Robertson *et al.,* 1982). Nevertheless, these three major information retrieval models have become the basic model of extending or conducting other information retrieval models such as the fuzzy set model and the natural language model.

Generally, a variety of techniques are also emphasised in locating information to help users to obtain information effectively and efficiently in a minimum time-frame. Chu (2003) lists five techniques under basic retrieval techniques and identifies four other techniques as advanced retrieval techniques. Boolean search, case sensitive search, truncation, proximity search and field search are listed as basic retrieval techniques. Fuzzy search, weighted search and query expansion are examples of advanced retrieval techniques that are used in some systems such as The Education Resource Information Center (ERIC) system. The Boolean search technique consists of keywords entered by the user with a combination of logical operators such as *AND*, *OR* and *NOT*. The *AND* operator is used when users want to narrow down the searching result while the *OR* operator is for broadening the search, whereas the *NOT*

operator is used to exclude unwanted keywords. These three operators were used only if more than one keyword was entered in the search box. However, the use of this Boolean search technique can vary from different search engines in which some search engines only accept the logical operator to be written in a symbol sign such as '+' and '–' rather than *AND, OR* and *NOT*. For instance search engines such as Google includes the *AND* operator when the user enters more than one words. Here is an example: when we search the keyword of *"cityrail timetable"* in the Google search engine, the results will return fewer results than the Yahoo search engine. This is due to the different search query algorithms applied. By using the Yahoo search engine, users need to add the plus sign (+) to instruct the search engine to include the words that have a plus sign in front of it. Otherwise, the query will automatically use the *OR* operator.

Case sensitive search is another retrieval technique that works by differentiating two different entities. For example, the word *Web* with the upper case of 'W' and the word *web* with the lower case 'w'. The difference is the word *Web* often refers to the application of the World Wide Web while the word *web* refers to spider webs. In addition, to be able to make the query to be specific, other techniques can be used especially to identify a term that has a specific letter as a basic word, which includes words ending with *-ed, -ing, -s,* and *-ion*. This stemming, striping or wildcard technique is known as truncation in which truncation techniques can be used to achieve different forms of a term, but actually referring to one common term or allocate words with alternate endings. For instance, we use the term *process** to search for documents containing terms of *processor, processed, processes* and *processing.* There are a variety of truncation symbols used by different search engines such as #, :, +, $, !, ?, and * (Large *et al.*, 1999). Proximity search is another search technique which specifies precisely how two terms should be or should not be included. This is based on the setting applied to the terms; for example, *information WITH retrieval* refers to two terms side-by-side which have to follow the sequence of the terms listed. Another basic technique that can be considered is field searching. This technique can narrow down the search results effectively, for example, by narrowing down the specific attributes such as an author search to specify the specific attributes of an author's name to narrow down the search.

A fuzzy search seems to be almost the same as truncation but it is designed to identify the data entry problem of a term misspelled. For example, if a user accidently typed 'computer' as *compter*, *compyter*, *coputre*, *cmputer* or *cmpter*, the system will detect the errors and propose suggested correct terms of *computer* when the fuzzy search technique is applied. A fuzzy query is useful for some kind of information such as linguistic, text statistics and data extraction where it can accept ungrammatical sentences or disjointed phrases, but with the requirement of using certain symbols such as + for *AND* (Baeza-Yates and Ribeiro-Neto, 1999). Ilhan and Duru (2006) describe a fuzzy query as a query that is closer to the human natural speaking language and intelligence. This type of query is referred to as fuzzy set theory that is classified under a fuzzy model or extended Boolean model (Kim and Lee, 2001). While there is continuous research on the fuzzy model, a different approach such as a knowledge-based approach (Chen and Wang, 1995) and extended fuzzy concept network (Chen and Horng, 1999) have been introduced and issued as a search technique. Furthermore, a fuzzy query can be performed in a more flexible and intelligent manner as this query can be conducted in a positive association, negative association, generalisation and specialisation queries as reported by Chen and Horng (1999). This extended fuzzy concept network is explained as fuzzy positive association, which relates to similar meaning; fuzzy negative association which relates to complementary; incompatible or antonyms concepts; fuzzy generalisation if the concept is included in an analytic or partitive sense; and fuzzy specialisation as inverse of its relationship.

2.1.2 The Effectiveness of Information Retrieval

Ranking can provide an indication of the document's relevancy based on queries, weighting terms become one of the preferred methods applied in some of the search engines. Based on the number of terms appearing in a document, the highest number of terms found, the highest ranking is assigned in the list and thus the higher the weighting is applied. Large *et al.* (1999) states a commonly used term of weighting mechanism is by including the value of the frequency of terms seen in the database, as well as in individual documents. In addition, a weighted search can be used to help users to obtain documents they need effectively. Weights are used to assign each search term to specify the relevancy of terms in documents. If two terms are given in

a query, the number of times these terms appear in each document will be recorded. For instance, term 1 occurs 3 times in document 1 and 4 times in document 2, while term 2 occurs 5 times in document 1 and 4 times in document 2. Thus, document 1 holds a weight of 7 and document 2 holds a weight of 9. The ranked output will be document 2 followed by document 1.

Query expansion is another technique that helps searchers to improve the search results when the first attempt was unsuccessful. This can be done by revising the search statement based on the first results. In other words, query expansion is similar to query reformulation that allows users to revise the query of retrieved results. Another search technique that has been acknowledged is the nearness query. This technique identifies the near meaning of some documents to each other from a huge database of documents available on the Web. This way can identify more relevant results that meet user's information needs when searching for documents. Colomb (2002) explains that nearness query language is different from the Boolean query in their search performance. Nearness query is more simplistic than the Boolean query, while the Boolean query allows searchers to control what will appear in its search result list by referring to the amount of information provided.

The effectiveness of information retrieval is not only based on how well the searcher makes a query and finishes the search task, but also how documents are being represented in the system (Baeza-Yates and Ribeiro-Neto, 1999). The searchers are responsible to express their need in the form of query language to be understood by the system. A logical view of the document is another aspect in the information retrieval system that needs to be well represented. Stored documents are frequently represented by a set of terms indexed or keywords that can be automatically retrieved. Representative keywords can also be automatically derived from the source documents either by automatic index extraction or generated by specialists. Chowdhury (2004) points out two different parameters for performance measurements: effectiveness and efficiency. In an information retrieval, the system effectiveness is measured by how far it can achieve its goals and fulfil the objectives while efficiency is how economically the system functions effectively. These include

factors such as user effort, response time and financial expenditure, which may include labour costs, databases, telecommunication and equipment costs.

The simplest and basic technique to evaluate effectiveness of information retrieval has been investigated through human effort called the "binary measure". Unfortunately, this technique of evaluation has experienced the difficulty of ill-defined terms by having different perspectives of human interpretation (Korfhage, 1997). This aspect of performance measurement is now being evaluated using precision and recall measures. Recall is a measure of searching effectiveness when it can retrieve the amount of relevant documents from the total of relevant documents that have been traced in the database. While on the other hand, precision is to access records that are totally traceable as relevant documents and to avoid receiving irrelevant documents in producing a high precision (Large *et al.*, 1999). Thus, performances can be expressed as ratio percentages, which are shown below:

$$\text{Recall ratio (\%)} = \frac{\text{Number of relevant documents retrieved}}{\text{Total number of relevant documents traced in the database}} \times 100$$

$$\text{Precision ratio (\%)} = \frac{\text{Number of relevant documents listed in the search results}}{\text{Total number of relevant documents retrieved}} \times 100$$

The ratio of recall can be higher in some cases but it is not possible to achieve 100% of recall and precision. Consider a situation where there are 100 documents in a collection, the search result returns 60 documents. This shows that there is 60% recall of the retrieved documents. While precision in a given situation where the search has retrieved 80 documents; 50 documents are relevant and 30 documents are irrelevant, then the precision is 62.5%. In these cases, the recall is said as having a high recall while precision can be said as having a high precision. Chowdhury (2004) reports that high precision would have a tendency to save users' time and effort. There is also a pattern of ratio of relationship between recall and precision. For instance, when searching is done in common or broad search terms, it results in high recall but low precision. On the other hand, searching for specific terms tends to have

low recall but high precision. Generally, users are not getting a high recall but may be pleased with a moderate level of recall.

Researchers have also identified different types of performance measurements to overlook the disadvantages of using precision and recall. Korfhage (1997) identifies a weakness of recall and precision as effective system measurements which includes the lack of information that can describe the relevancy of retrieved documents. Chowdhury (2004) proposes measures called 'fallout' and 'generality' that are closely related to recall and precision. Fallout is apropos to how the system can encounter the number of non-relevant items returned in a search result while generality focuses on the number of relevant documents received by the searcher. Zhou and Yao (2009) reviewed some evaluation methods which include Spearman's rank correlation coefficient, Kendall tau rank correlation coefficient and generalisations of average precision and concluded that the normalised distance performance measure is useful to identify highly relevant documents. Normalised distance performance is measured by taking the distance between the user and system ranking to determine the document's relevancy to implicate the number of agreeing, contradictory, and compatibility between document pairs (Yao, 1995).

Information retrieval evaluation or performance is not limited to precision, recall, fallout or generality measures. Broadbent *et al.* (2006) note in their report that there are several other measurement performances that can be categorised into: objective, system relative and subjective measures. Objective measure aims to obtain physical record as measurement such as speed and size of information retrieval process. Six aspects of measurement can be used which include index time, update time, query time, CPU utilisation, memory utilisation and disk space. System relative measures require a more time-consuming process of field experts to compare actual and ideal results with a system's answer set. These include calculations on precision, recall, precision histogram, mean-average precision, F-Measure and E-Measure. F-Measure is a measure of a test's accuracy, which considers the average rate in precision and recall and E-Measure focuses on a single measurement of information retrieval where the weighted harmonics mean of precision and recall is identified. Finally, subjective measure assesses user experience in retrieving information. This aspect of

measurement takes account of coverage, novelty, relative recall, recall effort, normalised recall, satisfaction and frustration. For instance, relative recall identifies the ratio of retrieved relevant objects to expected relevant objects.

2.2 Information and Database Search Problems

Based on literature review on information and database retrieval, the main problems of information and database searching can be categorised into five broad categories: (1) query formulation, (2) lack of user perseverance, (3) terminology and search terms, (4) lack of user experience, and (5) synonyms and homonyms.

2.2.1 Query Formulation

The aim of information searching is to obtain the most relevant documents to be retrieved. Searchers have to construct a query to be processed by the computer to retrieve useful information. Therefore, to achieve what we are searching for, we need to find ways of how the system can understand the user's query. Generally, searchers have at least a topic area or idea of what they want and then they transform their needs into a query form. Sometimes the first query will help to assemble a better query to gain better results. However, searchers usually make ad hoc queries once they know what they are looking for, and may become overwhelmed with huge amounts of information available (ter Hofstede *et al.*, 1996). This problem is described as "not knowing what they are looking for but they will know it once they have seen it".

Becker (2003) explains that students have trouble locating information they need, as they are facing difficulties as inexperienced searchers in which they do not really understand information organisation and retrieval concepts. In other words, students are unable to formulate the search query effectively. Jansen and Spink (2004) find that most of the users of major Web search engines used only one term rather than advanced operators. This shows that the users are not familiar with using advanced queries and have problems in formulating search terms with query operators. For example, Spink *et al.* (2001) report that 20.8% of users submitted two queries in one session of the Excite search engine log and approximately 31% of users submitted

three or more queries. Pass *et al.* (2006) shows that 28% of all queries entered are modified from previous queries.

2.2.2 Lack of User Perseverance

Beall (2007) applies the term ‘search fatigue’ to refer to frustrated searchers who are unable to find the desired information and often experience having search results that are unmatched to what they are looking for. For example, a searcher spent two hours sitting in front of the computer looking for the information in their project unaware that the information does not appear in the search database, even though the searcher may have tried many different ways to change the search terms.

Antell and Jie (2008) report that searchers tend to give up quite early after receiving several unsuccessful topic searches. While Chen and Rieh (2009) found that students are influenced by the topic familiarity when asked to do the search. Easy topics will make the search query easy and requires less time while harder topics require extra time to complete the search task. This is due to the fact that the terms of easy topics are terms that the searchers already had in mind. On the other hand, information that they are not familiar with will take the searcher a longer time to identify the exact terms to use or the best term that can be used to describe the query.

2.2.3 Terminology and Search Terms

Kagolovsky and Moehr (2003) emphasise two terminological difficulties in information retrieval which are related to semantic problems and a complexity of field of study. The semantic problem refers to how people can make mistakes using the same terms when referring to different terms and entities. For example, the terms ‘information retrieval’ and ‘information retrieval system’ refers to different meanings and applications although these terms look almost the same. Morupisi and Mooko (2006) observe that the most common errors made by the searchers are misspelling and sometimes they forget to include a space between keywords.

Kress *et al.* (2011) also reports some of the problems in terminology and search skills made by the searchers. Examples include searchers who fail to get the right starting point to search, searchers going back to the previous screen when there is no matched

results and searchers making spelling errors. Hölscher and Strube (2000) report that users with low domain knowledge notably made longer queries than high domain knowledge users. This is explained by knowledgeable searchers who know the exact term to use in queries. In terms of query formatting, Hölscher and Strube (2000) report that high domain knowledge users use less query formats in a search query, but they are more creative in using a variety of terms when searching.

2.2.4 Lack of User Experience

Dinet *et al.* (2004) report that novice searchers at the university level are not familiar with Boolean operators and they do not include as many operators as intermediate searchers (such as students who take the librarian course who have been taught how to use the Boolean operators) when making queries. This is due to the fact that students do not know how to manage the correct rules in a Boolean search. Al-Maskari and Sanderson (2011) discuss the results of their study which indicated that inexperienced students take longer than experienced students to obtain relevant search results. Hölscher and Strube (2000) discover that novice searchers are too rigid with one search engine and do not change in using the search approaches when they move to use another search engine. In addition, Antell and Jie (2008) report that students are unaware of basic search principles. For example, when students using two keywords in their first search attempt, which gave zero results, will later modify the search statement from using two keywords into three keywords which they hope can expand the search results, but unfortunately it will not, because the use of three keywords actually narrows the search result.

Chen *et al.* (1999) report that the sense of control of performing activities on the Web generally is influenced by how the Web environment is. For example, users find it easy to use the Web search engine if they are familiar with the environment and have the sense of being able to control the Web environment. Secondly, users go for deep search when their first search leads them to more information and they know how to go to the next step of research. Xie (2004) investigates problems of online databases and Web search engines and found that when categories of presentation are not well structured, irrelevant results are obtained. This happens when the searchers cannot provide the right picture or any clue of categories in which keywords are in.

Griffiths and Brophy (2005) explain that successful information search often relies on having some check-point items to support results. This means there should be more than one keyword to describe the right information they are searching for.

2.2.5 Synonyms and Homonyms

Beall and Kafadar (2008) describe synonym problems in full-text searches which result in some websites being missed depending on how common the word is used compared to the synonym of the word. For example, the words ‘mum’, ‘mom’ and ‘mummy’ all refer to the same meaning of ‘mother’. It is important to consider a range of words when using open keywords. Variant spelling is also another problem; for instance, spelling variant between American and British words such as color/colour. This kind of problem can be overcome by adding both words when querying. However, this approach can become complex when more than one word is involved.

Henzinger (2007) states that synonym problems happen when searchers come with an expectation that their query terms always appear in the search result. Thus, alternate query terms should be used in addition to the first synonym word. For example, the term such as ‘deadline’ could also be searched using other terms such as ‘cut-off period’, ‘due date’, ‘timeframe’ or ‘time limit’.

Another example of synonym problem is the use of short-form terms. Short-form terms make it easier to refer to, as in business names, centre names and familiar terms such as WWW for World Wide Web, USA for United States of America, and UOW for University of Wollongong. As the short-form term is often represented by the first letters in the sentence names, the possibilities of having the same short-form term cannot be discounted. For example, COM may be a shortened form for “communication” or “commercial” or even can be a short-form term for “Center of Momentum”.

Kohler *et al.* (2006) propose a system to overcome the homonyms problems for example, when we cannot differentiate the word “mouse” for an animal and the word “mouse” as a computer pointing device. This step is known as ontological indexing

in which the system uses ontologies and Natural Language Processing (NLP) to index text. This process supports word sense disambiguation by indexing text into concepts of ontologies. Other examples are when referring to the word “type” or “train”. The first meaning for “type” is “to write via keyboard” and on the other hand, it can be understood as “a sort of”. In the case of the word “train”, when people search for the word “train” it means “loco and trucks” or can be referred to as ‘to teach”.

2.3 Cross Language and Multilingual Information Retrieval

There are many languages around the world and the needs to search for information in languages other than English become important (Chu, 2003). There are two common systems that focus on the search language and content; they are Cross-Lingual Information Retrieval (CLIR) and Multilingual Information Retrieval (MLIR). CLIR is a search system where the target information listed in the search result is different from the language uses in the query search box, while MLIR is to have one language in a query search box and the target information contains more than one language. Thus, both the MLIR and CLIR applications require a translation process of the language used (Lin and Mitamura, 2004); either the translation is to convert the current query language to the desired language of the retrieved documents or the translation of the documents themselves is based on the desired language the user requested. Since the cost of translating documents is more expensive than a single translation of a keyword search, many researchers of the CLIR and MLIR applications prefer to use a query translation approach rather than document translations (Lin and Mitamura, 2004). A query translation approach translates the keyword entered in the query by using machine-readable dictionaries (MRDs), which are then used to search for the closed results. Document translations require a lot of effort and it is expensive to translate the whole document and be able to be identified in the preferred language search.

Yilu *et al.* (2005) identify three query translation approaches which translate queries into each target language before matching and retrieving related documents. These three approaches that can be applied into CLIR and MLIR systems are: 1) machine

translation based, 2) parallel corpus, and 3) bilingual dictionary. Lin and Mitamura (2004) stated in their paper that by using multiple machine translation based systems and the question sentence, translation correctness of the keyword term can be improved. There are some other researches also trying to improve the features of CLIR (Kapetanios *et al.*, 2008, Wang *et al.*, 2008b) and MLIR (Yilu *et al.*, 2005, Lin and Mitamura, 2004). Kim *et al.*, (1999) use a semantic category tree and collocation technique to improve precision of the CLIR system. A semantic category tree is proposed as pre-defined semantic categories, which apply semantic codes for the ambiguous words, which consider the equivalent translation of the word selection. For example, the keyword 'jumlah' in the Malay language is translated into 'amount' in the English language under the food category. However, the searcher may actually wish to refer to the word in the shipment category. In this case, the object is changed to the keyword 'volume'. By using this kind of semantic category tree, the concept can be created separately meaning there is an object known as 'amount' under the food category and another object for 'amount' under the shipment category. In addition, the collocation process solves the ambiguity in query translation by cross-checking with Electronic Dictionary Research (EDR).

Savoy (2003) and Abusalah *et al.* (2009) analyse and evaluate the CLIR system in obtaining the user's response and feedback to identify system potentiality in solving problems in cross languange retrieval. Savoy (2003) finds that retrieval effectiveness may be significantly reached by combining a bilingual dictionary and machine translation approach. While Abusalah *et al.* (2009) justify the effectiveness of an ontology-based approach, the CLIR system received higher precision ranking than the dictionary-based approach.

2.4 Overview of the OPAC System in an Academic Library

The Online Public Access Catalogue (OPAC) recently became one of the most convenient tools for knowledge searching. History of the OPAC system has been discussed by Large and Beheshti (1997). Chu (2003) points out five exclusive features of the OPAC systems. First, OPACs have been developed to keep the bibliographic information of all the resources managed by the library. Second, the

OPAC idea is an extension of MAchine-Readable Cataloging (MARC) records that have rules and standards to be monitored. Third, the OPAC system provides different types of searches such as the field search, keyword search and Boolean search. Fourth, OPAC has access to CD-ROMs and online databases. Fifth, the OPAC system allows users to browse the library's records that can be accessed by authors and titles. In the OPAC system, the use of Z39.50 protocol enables the search process to retrieve information from different types of sources, particularly online systems such as ACM, IEEExplore, Proquest, Springerlink and Science Direct. The standard Z39.50 is a protocol that specifies the data structures and interchange rules that allows communication between the client and the server (Lynch, 1997). There are five methodologies of data collection that can be used to help in OPAC usage analysis. These five methodologies are: experiments, interviews and questionnaires, observations, think aloud and transaction logs. The data collection is based on OPAC usage, searcher's behaviour, and activities to identify problems of relevance results returned. Furthermore, different OPAC systems provide different types of search techniques to help users find the information they need.

Lau and Goh (2006) propose three recommendations to improve the quality of the OPAC system, which include: (1) interactive query reformulation by recommending alternative query terms, (2) browsing selected items in the search result list through hyperlink records, and (3) content-sensitive assistance which can identify user's searching skills either as an expert or novice user, which in turn can provide support of terms and document recommendations. The relationship between users and the OPAC system can be identified using the Transactional Log Analysis (TLA). Transactional technique is used as a gathering tool to record the communication and interaction between humans and the system. TLA can be used to determine the most frequent types of search queries used (Jones *et al.*, 2000, Villén-Rueda *et al.*, 2007) and to overcome search difficulties (Bangalore *et al.*, 1998). Villén-Rueda *et al.* (2007) report that queries by title returns a higher percentage than using queries by subject heading or queries by authors. Meanwhile Yu and Young (2004) discover that query by subject heading requires the searcher to have knowledge in subject titles or author's information to start the search.

The author has investigated an OPAC system offered by the Public University's Library in Malaysia. The Ministry of Higher Education (MOHE) in Malaysia listed 20 universities in which each of these libraries have developed their own databases and library catalogue systems. These online library catalogue systems offer intra-access as well as inter-access. The OPAC systems provide different types of query searches for different libraries. For example, the Web OPAC for Universiti Teknologi Malaysia (UTM) offers open keyword searches with a filter function (such as author, title, thesis and journal). However, in other OPAC systems, the system may offer a match term option (such as exact or partial keyword). In another example, the OPAC system for Universiti Malaya (UM) offers keyword searches or browse for options such as word or phrase, author, title, subject or periodical title.

Different languages used in the search process can result in major differences in the number of results returned. Most of the OPAC systems in the public universities in Malaysia have been developed using the English language. Some systems provide bilingual interfaces such as the system used by Universiti Kebangsaan Malaysia (UKM). The bilingual system interface unfortunately does not cater for bilingual search results. On the other hand, only keyword matching is performed and a query conducted in English terms do not return results in both the English and Bahasa Malaysia terms. The author has conducted a short experiment on the OPAC system in two universities: Universiti Malaya (UM) http://www.pendeta.um.edu.my/uhtbin/cgisirsi/x/x/0/49/ and Universiti Kebangsaan Malaysia (UKM) http://gemilang.ukm.my/ to investigate search results return using different search queries and language.

2.4.1 Pendeta WebPAC: UM Library Public Access Catalog

In the case of UM online library catalogue, it offers a quick search option either by keyword or browse. The search can be conducted by *phrase*, *author*, *title*, *subject* or *periodical title* as shown in Figure 2.1.

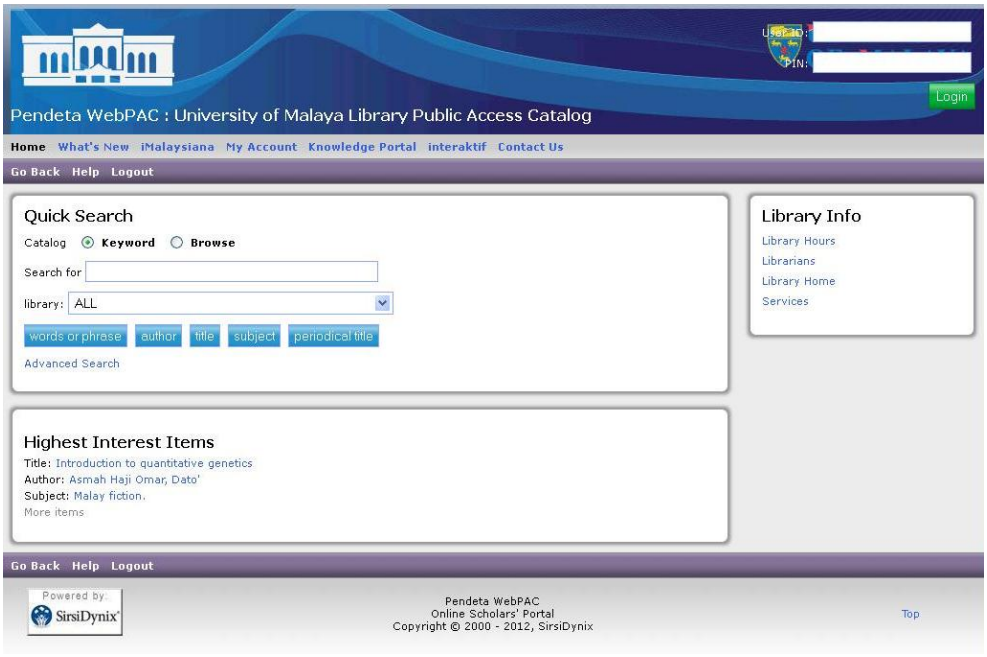


Figure 2.1: Pendeta WebPAC System for Universiti Malaya (UM)

The author has run a sample query to search by using "computer" as a keyword in the Computer Science Library (see Figure 2.2). Figure 2.3 displays a screen shot of the returned results. In this example, the keyword search processes the search by referring to only the Computer Science Library as the catalogue resource and the search is based on a title that consists of the keyword "computer". The result obtained shows that 11 titles that matched the keyword "computer" has been found.

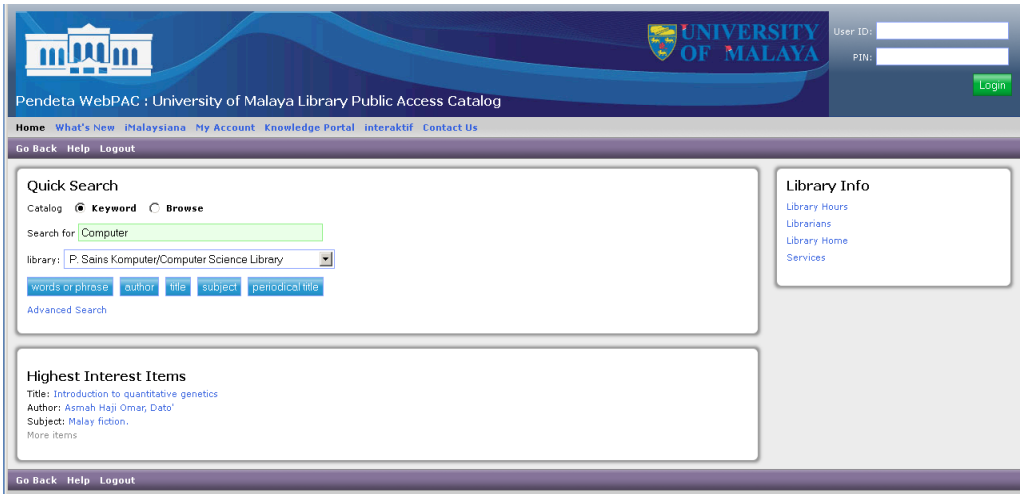


Figure 2.2: Screen shot of the query sample

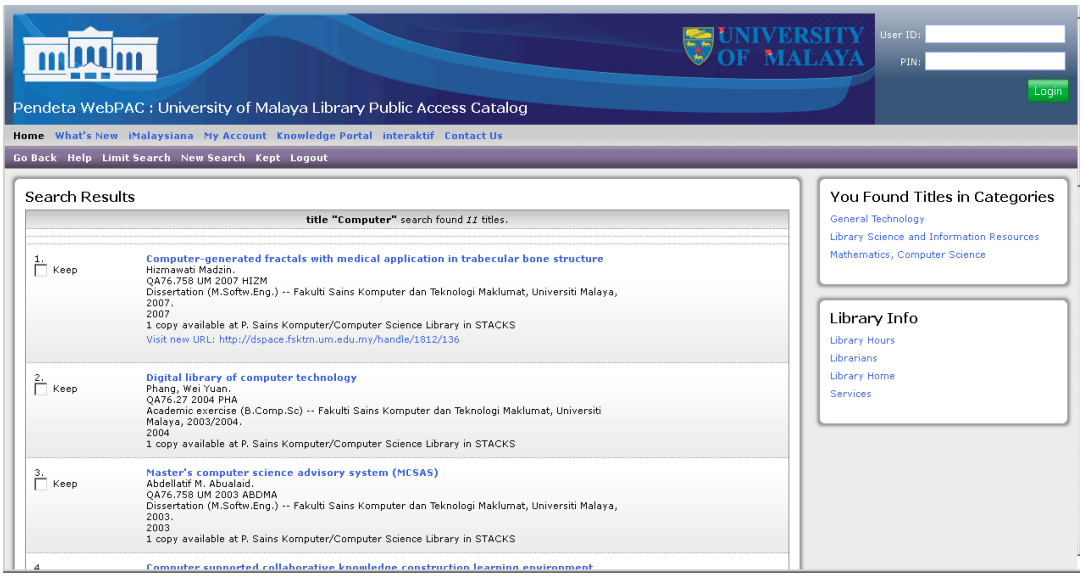


Figure 2.3: Screen shot of the query result

2.4.2 Gemilang: UKM Library Catalogue

In another experiment, a search was conducted using the UKM online library catalogue. There are two search options: browse and quick search (as shown in Figure 2.4). The browse search option allows the user to enter keywords or phrases for authors, titles, subjects, publishers, call numbers or journal titles. While a quick search offers four different types of sources, which are anywhere, books, newspapers or periodicals. The differences between these two search options can be described as follows. The browse search results returned only the preview option of showing the list of the subject (such as: author, title, subject etc). In this case, the user has to browse the list one-by-one to get to the right documents, and details are only shown when the user clicks the preferred item in the list. On the other hand, the quick search option previews details of results in a complete list.

Figure 2.4: Gemilang OPAC System of Universiti Kebangsaan Malaysia (UKM)

The author runs a sample query using the "Browse" option, with the keyword "Computer" entered in the search textbox and wishes to have the results returned for "title". Figure 2.5 shows a sample screen shot of the returned results. The result shows a list of subjects returned that matched the subject "Computer". For example: subject "computer" has 4 records under that subject category while subject "computer adaptive testing" has 3 records under that subject category. To view details of each subject the user can click on the subject listed and full results will be displayed.

Figure 2.5: Screen shot of the quer

The author has also conducted another simp performance of the OPAC system at UM and UKM. The query is to search for titles with the keyword 'computer' in both English and Bahasa Malaysia languages. Table 2.1 shows the comparison result. When the keyword 'computer' is entered (in English) a total of 24,464 records were retrieved from the UM system while the keyword 'komputer' (in Malay) returned only 1814 records. However, when the keyword is entered as 'computer *OR* komputer' the UM system only returned 25,178 records which are 1100 less than the combined total records of using 'computer' (24,464) and 'komputer' (1814). This can happen when there is a record with both the English and Bahasa Malaysia words included in one title (for example: a translation book). Another possible reason is the field other than title (for example:

category or subject field) were also included in the search field to find the match keyword to the requested query.

Table 2.1: Search results of English and Malay keywords entered in the OPAC system (as of 5th March 2012)

	University	
	UM	UKM
Computer	24 464	17 961
Komputer	1 814	1 827
computer OR komputer	25 178	1 104

2.5 The Semantic Web

According to Berners-Lee *et al.* (2001), the Semantic Web is an extended version of the current Web in which information available in the WWW are accessible and it increases the machine capability to analyse, assemble and filter Web documents. The machine capability refers to how machines can interact with one another through the use of software agents to generate search results (Berners-Lee *et al.*, 2001). Researchers believe that the Semantic Web not only has the ability to exchange data between machines, it will also bring meaning to the documents and is capable of understanding the documents to fulfil users' needs (Fensel, 2000, Java *et al.*, 2007). Hendler (2005) suggests that not only Web documents have meanings in the Semantic Web, the semantic technology can also help users in searching information in the best possible way to suit individual's needs.

An example of how the Semantic Web technology can be applied was described by Berners-Lee *et al.* (2001) to show how a software agent helps a user to arrange an appointment for their mother to meet a doctor, and to attend a series of physical therapy sessions. The resultant plan presented meets the requirements and needs instructed by the user. In this example the user only needs to instruct the personal agent and the agent responds to the task by searching for information in the Semantic Web. When describing the capability of what the Semantic Web can do, Berners-Lee

et al. (2001) explains that the agent is a main tool that crawls through the Web to perform the task such as searching. While locating for documents or information, the agent looks up several lists of providers and matches the providers with the preferred requirements of the user. Another Semantic Web scenario is a patient requesting for a doctor's prescription (Kim *et al.*, 2002). In this example, the patient only needs to instruct the agent to obtain the prescription. The agent then contacts the doctor's agent to check the records of the patient and issues a new prescription. Then, the patient's agent contacts the preferred pharmacist's agent to arrange a delivery after all the information (including financial data) has been provided, verified and processed. Even though the process includes private data such as credit card information and the user's personal data, the information that travels over the Web is protected by the trust and security layers of the Semantic Web.

In other applications such as online shopping, the agent is able to search and compare products such as product types, brands, prices and features (Davies *et al.*, 2004). Kim *et al.* (2006) implemented an online shopping mall employing an intelligent shopping agent. The agent engages an intelligent search engine for customers based on the Semantic Web and integer programming technologies. The integer programming method in this example solves the problems to achieve the goal the customer sets when buying a CD from an online shopping mall. It starts with the customer identifying the necessary keywords and the ontology provides the CDs information from all registered shopping malls. The result lists the available CDs with information such as price, delivery rates and discounts the shopping malls have at that time. In this way, the customer can purchase the CD with optimal results that matches the requirements of the customers in terms of product price, delivery charge and other requirements.

In the Semantic Web, the computer becomes an intelligent machine that is able to fulfil user's needs and restructures search queries and mechanisms to produce more accurate outputs. For example, the software agents are able to understand the meanings of the data, perform appropriate operations to ensure data and information meet the search criteria. In addition, the software agent will perform checks to ensure the legality of results or to derive a new construct of results using inference and

logical reasoning (Cardoso and Sheth, 2006). This, in essence, is how we envisage an intelligent machine should perform (Cardoso and Sheth, 2006, Gerber, 2006). The aim is to design computers or machines in the Semantic Web that are capable of performing tasks usually performed by humans, such as generating queries and searching for the right results. As described by Thuraisingham (2004), the Semantic Web is a very intelligent technology that assists humans to collect information from heterogeneous resources and databases. These capabilities are achievable due to the unique characteristics of the Semantic Web that include the use of metadata and logical inference, as evidenced in the layer cake diagram proposed by Berners-Lee *et al.* (1994) as shown in Figure 2.6.

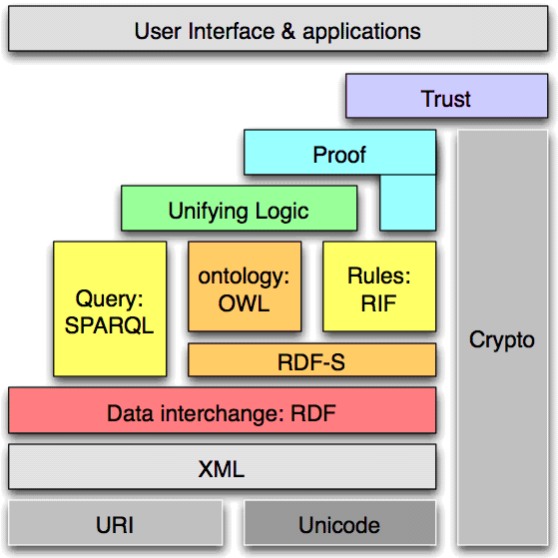


Figure 2.6: Semantic Web architecture (Source: Berners-Lee, 2001)

The architecture is layered with nine different types of layers and a layer of cryptography that resides alongside the other layers. The basic layers are URIs and Unicode. URIs is a platform of providing web-based resources by identifying the resources but not working as a set of directions to find the resource's location. This is followed by the second layer of Extensible Markup Language (XML) which is a language of metadata; subsequently the stack is upgraded to Resources Description Framework (RDF) language. Although the XML is a markup language that enhances the meaning of the data itself, RDF makes the data as "machine-processable" rather than "machine readable". This feature of "machine-processable" is referred to as the

ability of the machine to process metadata; provides interoperability of machine-understandable as to facilitate knowledge sharing and information exchanges between applications.

The next layer is RDF-S, which is Resources Description Framework Schema, which is used to describe the relationship between data, and helps the computer system to use, share and convert the terms easily. All the terms used in this layer can be presented using an ontology language. Web Ontology Language (OWL) is one of the ontology languages that can be used where the rules are applied to have a reasoning feature. The logic layer needs to be applied so it can analyse the capability of interoperability of data where the proof is to validate items and to provide a conclusion of data interpretation or assumptions. The web of trust is a final layer on the Semantic Web application where the source of the information provided can be trusted. Finally, the cryptography layer is the backbone of the Semantic Web architecture to provide a secure and protected transmission of data over the Internet using digital signatures.

Semantic is the meaning of the data while semantic interoperability is a kind of machine capability to intelligently understand the meaning of information that traverse across the Web. This must be done without human intervention in terms of understanding the meaning of data available to be shared between machines. The information must be accessible by the agent and then interpreting it to become a machine understanding format. For example, in human normal life, the semantic interoperability is similar to a scenario where a mother refers to the term “nursery” as a place for parents to send their children to kindergarten in a school or a playgroup. On the other hand, another person may think of it as the place for plant sales outlet or a garden centre. Sandra (1995) suggests that semantics discovery and representation, and management of mapping can help to address semantic interoperability problems with repository technology to manage complex mapping. Semantic discovery offers common understanding between the searcher and system provider of meaning of data.

In fact the Semantic Web has been ironically identified as a web of metadata. There are a number of mechanisms that can relate or map knowledge between existing ontologies. Maedche and Staab (2001) state that people intend to manipulate or extend an existing ontology rather than build a new one in line with the objective of ontology development to allow knowledge reuse. There are many ontologies that have been developed and produced which include Textpresso (http://www.textpresso.org/) and WordNet (http://wordnet.princeton.edu/). Textpresso helps in providing information based on ontology, which holds most of the content in biological literature, while WordNet is a lexical database offering concepts stored with hierarchical relations.

2.6 Ontology

Ontologies have been known as a database of terms that justified a domain to be used and shared in a global area (Borst, 1997). Ontology becomes a model of real word to represent a domain of knowledge. This new technology has been used in the Semantic Web although the original word of ontology is being borrowed from the philosophy discipline, which defines the concepts of things. Thomas (1993) explains the real definition of ontology is a systematic account of existence, however in computer science, ontology is a representation of precise specification to form a concept. Thus, ontology is described as formal specification of terms in the define domain and identifying any relations existing in between the terms. Ontology enables people or machines to retrieve the desired information with an understanding of the meaning of terms used in the domain and share common vocabularies used in the same domain (Wang *et al.*, 2008a). Therefore, the use of ontology is about using, reusing and sharing domain knowledge of terms concept. Many ontology classes have been developed recently and are kept in a database to be used or referred to by others as knowledge/resource sources. Ontologies are not only used in the field of Semantic Web but also in many others fields such as artificial intelligence, software engineering, biomedical informatics, library science, and information architecture.

There are two types of ontologies according to two dimensions of perception: the amount and type of structure of the conceptualisation and the subject of the conceptualisation. The first dimension, according to Heijst *et al*. (1995), includes: (i) terminological ontologies, (ii) information ontologies, and (iii) knowledge modelling ontologies; whereas the second dimension includes: (i) domain ontologies, (ii) generic ontologies, (iii) representation ontologies, and (iv) application ontologies. The first dimension with terminological ontologies is referred to as ontology that defines the terms to represent knowledge in the domain of discourse, such as medical or biological domains. Information ontologies are defined as records structure of a database, which is a flat structure, unlike the knowledge modelling ontologies, which have a richer structure of database, such as involving distinction and decision-making processes. To refer to the second dimension of ontologies, domain ontologies refer to specific particular area while generic ontologies refer to domain ontologies across many areas. Representation ontologies are supposed to be naturally present in general contrast to application ontologies, which are specifically designed to the particular application such as the Marine Metadata Interoperability Project (MMI) (https://marinemetadata.org/).

Holsapple and Joshi (2002) present five approaches to ontological design: (1) inspiration, (2) induction, (3) deduction, (4) synthesis, and (5) collaboration. Inspirational approach starts the design idea by collecting individual personal views and creativity to construct the domain context. Inductive approach is based on the observation and analysing of current or specific domains to apply to particular domains. Deductive approach adopts some general principles to construct a new domain while the synthetic approach applies some potential characterisation from the existing ontologies. With the collaborative approach, the approach relies on human participation, which involves individual reflection and viewpoints to get along with the collaborative process.

How these ontologies can be developed depends on how or what method is being used. Uschold and Gruninger (1996) conclude that there are five steps in the process of ontologies development: (i) identify purpose and scope, (ii) building the ontology, (iii) evaluation, (iv) documentation, and (v) guidelines for each phase. In the second

step of building ontology, it includes: (a) ontology capture, (b) ontology coding, and (c) integrating existing ontologies (Uschold and Gruninger, 1996). The first step in building the ontology is by considering when there is a clear idea on what ontology is going to build, and then the domain of the ontology can be set with purpose and scope of the domain identified earlier. This idea can then be extended to the second step of developing domain ontology by providing information of ontology capture, coding and with attention to consider using an existing ontology. The third step is important to identify whether the ontology is in a good form of classification and relationship in its domain to bring effectiveness of knowledge sharing. In the forth step, the idea of having documentation is to allow knowledge sharing by preparing the problems faced in existing ontology with the important assumption together with the concepts definition based on type and ontology purpose. In the last step, the initial guidelines are provided which consists of clarity, coherence and extensibility.

Some other methodologies for building ontology have also been discussed by Fernandez-Lopez *et al.* (1997); and Corcho *et al.* (2003a). Corcho *et al.* (2003a) have review and compare the main methodologies for building ontology such as METHONTOLOGY (Fernandez-Lopez *et al.*, 1997) and On-To-Knowledge methodology (Steffen *et al.*, 2001). Fernandez-Lopez *et al.* (1997) propose the ontology development process to start with planning, specifying, knowledge acquisition, conceptualising, formalising, integrating, implementing, evaluating, documenting and maintaining the process. This methodology is used in most ontology development processes (Lopez *et al.*, 1999, Brusa *et al.*, 2008) and has also been extended to allow collaborative edition of ontologies at the knowledge level (Arpírez *et al.*, 2001). On-To-Knowledge methodology takes into consideration the process of ontology development from the early stage of setting up the project until the final level of the application which consists of: feasibility study, ontology kickoff, refinement, evaluation and maintenance (Steffen *et al.*, 2001).

Table 2.2: Ontology Development Tools

	Ontology Development Tools	Main Features	Ontology Development Process	Strength	Weakness
1.	OilEd	Ontology editor for building ontologies using Ontology Interchange Language (OIL).	IOilEd knowledge model is based on description logics.	Allow users to exploit the full power of an expressive web ontology language (OIL/DAML+OIL). Reasoning is used to support the design and maintenance of ontologies.	No support for versioning or for working with multiple ontologies. The reasoning support provided (FaCT) is incomplete for OIL extended with concrete data-types and individuals, and does not include additional services such as explanation.
2.	OntoEdit	Ontology editor	Methodology-guided development of ontologies.	Ontology development with the help of inferencing. Extensibility through plug-in structure.	No support on built-in inference engine, DBMS, collaborative working and ontology library.
3.	WebODE	Integrated technological support for many activities of the ontology lifecycle, technological support for ontology development methodologies, ontology interoperability.	Provides a default form-based web user interface to create ontologies according to the knowledge model aforementioned.	Extensibility via plug-in. Graph view. Multi-user. Merging.	No support on ontology library.
4.	WebOnto	A tool for collaboratively browsing and editing ontologies, which uses a Java client to alleviate the interface problems generated by HTML.	Provides direct manipulation interface using graphical representations to present ontology constructs.	Incorporates the sketching and synchronous communication tools. Provides a direct manipulation interface using graphical representations to present ontology constructs.	No support on extensibility and merging.
5.	Protégé	An open-source tool that allows	Protégé knowledge model	Protégé 2000 provides ontology	Does not provide real support for

		developers to create and to manage terminologies and ontologies.	is frame-based.	editing functionality on different levels. It is possible to query the ontology. Plug-ins are provided for querying based on F-Logic, merging and annotation of the ontologies with WordNet.	multiple users. There is no support for multiple changes on the same component. Users are not notified about changes made by others. The tool does not allow two classes or attributes with the same name.
6.	OntoSaurus	Ontology browser.	Linking domain specific terms to an existing ontology and extending it.	Share knowledge across systems; use object-oriented Common-Lisp-based as a web server and data knowledge can be browsed and edited on any platform.	Use HTML for the client interface that generates interface design problems. It does not support ontology development from scratch.
7.	Ontolingua	Provides a suite of ontology authoring tools and a library of modular re-useable ontologies.		Enables renaming of non-logical symbols from multiple component ontologies and that disambiguate symbol references during input and output.	No support for graph view and extensibility.
8.	KOAN	Builds on available resources and provides tools for the engineering, discovery, management, and presentation of ontologies and metadata.	Includes a multi-lingual Ontology Engineering and Evolution Environment (OntoMatSOEP) that allows the manual development and maintenance of ontologies.	Provides objects representing various pieces of ontology, such as Concept, Relation, Instance, Objects for creating and applying changes to ontology entities as well as objects providing query facilities.	No support on graph view and merging.
9.	SymOntoX	Ontology management system that is capable of developing and managing several ontologies (business and enterprise).		Support collaborative and distributed ontology authoring activities.	
10.	Retrievalware 8.0	Is an enterprise search engine emphasising natural language processing and semantic	Combines Convera's proven enterprise search and categorisation capabilities with a new dynamic	Uses a semantic network to expand queries for more complete recall. Profiling feature	

		networks.	classification methodology to help organisations automate knowledge management and discovery processes.	filters content to enable real-time monitoring of information in live data sources.	

Other than ontology methodology, there are currently many tools available to develop ontology. Some of them are: OilEd (Bechhofer *et al.*, 2001), OntoEdit (Sure *et al.*, 2002), WebODE (Corcho *et al.*, 2003b), WebOnto (Domingue, 1998), Protégé (Rubin *et al.*, 2007, Tudorache *et al.*, 2008, Corcho *et al.*, 2003b, Corcho and Gómez-Pérez, 2004), OntoSaurus (Swartout *et al.*, 1997), Ontolingua (Thomas, 1993, Farquhar *et al.*, 1997b), KAON (Bozsak *et al.*, 2002) and SymOntoX (Missikoff and Taglino, 2003). For example, OntoSaurus (Swartout *et al.*, 1997) presents ontology browsers and editors to support a collaborative vision of ontology development. Farquhar *et al.* (1997a) develop tools to allow ontology sharing and at the same time provide services to publish, browse, create and ontologies editing stored on their Ontolingua server. See Table 2.2 for detailed comparisons of ontology development tools.

While there are many ontologies that have been developed and produced formerly as discussed in the previous paragraph, the feature of sharing and reusing the existing ontologies is concerned as ontology development is deemed to be time consuming if it involves constructing new knowledge bases from scratch and it is labour intensive. However, existing ontologies can be imported and mapped or merged into a new system in order to be re-used. Mapping and merging are two different concepts in ontology. Mapping concerns trying to bring the concept from the first ontology (source ontology) into the second ontology (target ontology); whereas merging concerns building a combination of ontology from the existing ones. Noy (2004) differentiates mapping and merging as follows: mapping involves two separate ontologies in which the first ontology is kept as a principle and the other part of ontology (second ontology) is related to consistency with each other through class or subclass mapping (between ontologies); while merging imports all related classes in

one ontology. For ontology development, we can choose either to build ontology from existing ontology by merging from two or more ontologies or we can build the ontology by mapping the source ontology to new concepts in a new ontology. All these processes can be done manually but as the ontology becomes larger and complicated, it needs automatic tools or at least semi-automatic tools to deliver this transformation efficiently and conveniently.

The mapping process can be divided into two categories which are: manual mapping and automatic (semi) mapping (Yu and Peng, 2007). The manual mapping is likely to be a step-by-step approach to be followed. The main problem with manual mapping is when the ontology becomes large, resulting in larger classes/entities of databases which makes it difficult to maintain. In contrast, automatic ontology mapping is an effective alternative, which is categorised into string matching and text classification techniques. This approach can be done in a very short period, rather than manual mapping with a very high degree of data accuracy, to fulfil the growth of data available in the Web.

Kohler *et al.* (2006) explain the steps of how to use several mapped RDF ontologies for concept-based text indexing. The steps are: (1) importing and mapping ontologies; (2) spidering; (3) indexing; (4) ontological indexing; and (5) front end. The mapping process can be achieved by importing different ontologies using a RDF-parser and mapped with the equivalent concepts of different ontologies. Before the indexing step, the web crawlers are used to gather information from the Internet by searching and downloading ontologies. Then, the indexing process is done word-by-word using previous information. Finally, ontological indexing is semantically set to the match ontological concepts to avoid word sense disambiguation. After all, users are able to search based on keywords or ontological concepts using front-end servers.

In the scope of ontology merging, when a system is integrating two or more ontologies to be merged in the same domain, this technique can offer more options of data availability for knowledge sharing. However, it can give rise to another problem: how the ontology can be integrated to solve interoperability problems. Yu

and Peng (2007) propose an approach that can semantically match concepts defined in one ontology with similar concepts in another ontology. This approach considers getting example patterns from the Internet and calculates the probability of two concepts similar to each other to find a code pattern of the similarity. Yu and Peng (2007) used automatic text exemplar software called Rainbow (McCallum, 1996) to retrieve word exemplars from the Web, to be used in text classifications in which the probability of relevant concepts is calculated automatically and used in the mapping tool.

Another method proposed by Swartout *et al.* (1997) is the use of merging tools that applies the merging process of WordNet with the English dictionary, which uses two techniques to analyse and recognise the match terms or concepts. The first technique looks for similarities of the textual definition, followed by the second technique to identify associations between unambiguous terms in both ontologies in the hierarchical structure of the ontologies. For example, the individual of 'cabin baggage' in one ontology and the individual of 'hand luggage' in another ontology are both listed under 'aircraft/luggage/size' classes, hence the tool will propose these terms as having a possible association. Table 2.3 shows a comparison between ontology mapping and ontology merging.

Table 2.3: Comparison of ontology mapping and merging

No.	Themes	Merging	Mapping
1.	Definition	Building a combination of ontology from the existing one.	Bring the concept from the first ontology (source ontology) into the second ontology (target ontology).
2.	Tools	iPROMT (Noy and Musen, 2003), Chimaera (McGuinness *et al.*, 2000), and OntoMerge (Dou *et al.*, 2005).	ANCHORPROMPT (Noy and Musen, 2003), GLUE (Doan *et al.*, 2002), and OBSERVER (Mena *et al.*, 2000).
3.	Method	FCA-Merge (Stumme and Maedche, 2001).	IF-Map (Kalfoglou and Schorlemmer, 2003).

Web Ontology Language (OWL) is one of a web ontology language, which has become W3C recommendation to build a repository of knowledge. It has been created and intended for showing a relationship between classes, properties and instances. The ontology will be used as a real model of the real world that has been transformed to a simulated scenario to represent a domain of knowledge. Ontology itself is used to describe a domain, have its main objectives to let data on the Web to be understood and shared by the web agents. OWL is designed for the web agents to process the content of the information that is represented by the ontology terms and intelligently communicates with other agents in terms of information sharing. OWL is the most mature language for modelling ontologies (Lange *et al.*, 2007). It has become a powerful language as this language is able to represent machine interpretable content rather than only can understand the meaning and semantic of the documents on the Web. OWL has three sublanguages designed to provide more options and functions specifically for user which are: OWL Lite, OWL DL, and OWL Full. The choice between OWL Lite, OWL DL and OWL Full depends on how big the ontology is that we want to build and how complicated the ontology is.

Semantic Web technologies can be used to overcome problems in information retrieval systems. d'Amato *et al.* (2010) attempt to improve the resource retrieval process by exploiting tree-indexes in which a new concept of clustering algorithm is applied to available resource description by using description logic languages. Chiang *et al.* (2001) propose a full semantic search system known as the Semantic Web Query (SWQ) which uses context ontologies to perform semantic search filters. SWQ focuses on web page searches on the Internet to formulate appropriate web queries through domain semantics by refining user queries through a set of keywords and filtering the domain ontologies in the context ontologies. SWQ is detailed in seven steps of query processing and page retrieving: (1) web query parse; (2) ontology determination; (3) synonym determination; (4) web query formulation; (5) determination of web page relevence; (6) filter search; and (7) page ranking. The first step in SWQ requires keywords, some Boolean operators and some search parameters to enable a right selection determined by the user to the context ontologies in the second step. In step 3, the keyword synonym is given which is then passed to the search engine in step 4 to search for the right context in the right type

of data search engine, such as graphic or video formats. In step 5, the most relevant web pages will be returned with the URL link and the text fragment of the web page is matched and ranked to produce a list of reduced relevant web pages. Step 6 is where those reduced relevant web pages are passed through the filter search of readability, document structure and word sense. Finally, in step 7 the web pages are ranked after comparing the numbers of relevant web pages to the number of web pages requested by the user.

In addition, Chiang *et al.* (2001) apply semantic search filters to rank the relevant web pages by employing three search filters to increase precision of web searches: (1) readibility filter, (2) document structure (layout) filter, and (3) word sense filter. The readibility filter uses the Flesch-Kinkaid readibility score in which the score is computed based on the number of syllables per 100 words and the number of words in sentences to identify relevant web pages. However, the second filter option which is a document structure filter, is not often used to rank relevant web pages as the lack of a term in a document structure will affect the page being ignored to return as the relevant web pages. For instance, a dental clinic is represented by the terms {clinic, diagnosis, treatment, anaesthetic, dentist} and a medical clinic is represents by {clinic, diagnosis, treatment, anaesthetic, general practitioner}. However, when in the search process a web page is identified as lack of term "general practicioner" and "treatment" in their index terms, the page will not be presented as a medical clinic layout meaning the web page is inappropriate for the given requested layout. In the third filter of word sense, it parses the text snippets from a web page and then compares to the *Word_Sense* property in the ontology terms to identify the correct word sense. For example, the word sense filter will reject a web page that contains the phrase 'They have to record the show' because the term 'record' is not used as a noun which is not appropriate to the definition of the term selected in the search request for 'Athletes world record'.

Rad *et al.* (2010) propose a new approach to retrieve information different from user language by the following processes: (1) extension of English query phrases, (2) phrase translation, and (3) use semantic disambiguation by calculating the translation probability. In the first process, extension of query is made before the translation in

which this process is identified and can lead to good query. This process extracts all synsets of any entry for the keyword by referring to the WordNet database. The second process is using a dictionary-based method, particularly bilingual dictionaries for phrase translations instead of word-by-word translations. In the third process, the calculation of combinational translation probability will be applied which needs an extract of different meanings in the English query by referring to a bilingual dictionary and by using a textual corpus. Kapetanios *et al.* (2008) propose a different approach, which is a parametric linguistics-based approach in which the linguistic knowledge structure is instilled to the submitted query, thus increases the precision of the semantic translation query and solving the translation ambiguity problem.

2.7 Conclusion

Issues of database searching and retieval methods have been presented in this chapter. The issues identified show that current approaches of database search design requires users to have a certain level of skill in search techniques, such as knowing how to use appropriate keywords and Boolean operators to search. Novice researchers often have difficulties in conducting searches using traditional database systems due to a lack of search skills. The novice researchers find it difficult to formulate a query. They find it difficult to transfer what they have in mind to correct search query formats. In addition, the search results can sometimes be unmatched to what users have expected, retrieving too many results or nothing at all. This problem is compounded with terminology problems as some users find it difficult to use the same keywords as what is indexed in the database. They also face problems when they need to use keywords that are different from their native language. Very often, the system cannot conduct bilingual searches. Some systems use bilingual searches consisting of cross-language and multi-language information retrieval to overcome the problem. Major problems related to search queries can be categorised into: (1) query formulation, (2) lack of user perseverance, (3) terminology and search terms, (4) lack of experience in searching, and (5) synonyms and homonyms.

Recently, Semantic Web technology has been used in information retrieval. In this chapter, issues relating to Semantic Web architecture as well as the capability of

Semantic Web technology in understanding the meaning of information across the Web have been presented. Ontology which is used to represent a domain of knowledge to allow use, reuse and sharing have also been described. Knowledge sharing through ontology data-store can be achieved by creating the ontology either from scratch, mapping into similar domains or merging between related domains. Ontology can be used to define the meaning of data and relationships between resources. An ontology-based search approach can provide an opportunity to overcome problems related to search queries to be presented in the next chapter.

CHAPTER 3: DEVELOPMENT OF AN ONTOLOGY-BASED SEARCH FRAMEWORK

This chapter proposes an ontology-based search framework that can assist novice users in query formulation. Major information retrieval issues faced by novice users are presented in this chapter. The major issues identified are: (1) query formulation, (2) lack of user perseverance, (3) terminology and search terms, (4) lack of user experience, and (5) synonyms and homonyms. The chapter discusses the development of an ontology-based search to assist novice users in query formulation. Application of the proposed framework is illustrated with an example.

The chapter is organised into two sections. Section 3.1 summarises major retrieval issues faced by novice researchers. Section 3.2 discusses the development of ontology-based search approach. A conclusion is presented in Section 3.3.

3.1 Major Retrieval Issues

Online database systems have been used widely by researchers when conducting research for research topics. However, to be able to process queries efficiently is not an easy task. Common problems encountered in full text searches include lack of precision and relevance. Beall (2008) explains the weakness of full text searches and can include synonym and homonym, variant spellings, short form of terms and different languages or dialects used. As presented in Chapter 2, these problems can be categorised into five groups: (1) query formulation, (2) lack of user perseverance, (3) terminology and search terms, (4) lack of user experience, and (5) synonyms and homonyms. This section reiterates these five categories by focusing on problems faced by novice researchers in the search process.

1. Query Formulation

Becker (2003) explains that students have trouble locating information they need as they are facing difficulties as inexperienced searchers, in which they do not really understand information organisation and retrieval concepts. In other words, students

are unable to formulate the search query effectively. Very often human interpretation is required in determining which documents are relevant and which are not. In some situations, students constantly change the search terms during the search process. For example, students may use the terms of "opening hours", "operating hours" or "store hours" in different attempts to search for information for "store shopping hours".

2. Lack of User Perseverance

Chen and Rieh (2009) conclude that time-related issues can affect search results. Searchers sometimes find that simple searches can consume a longer time than initially planned. In addition, searchers who have a low level of information seeking skills may use generic words to search, which can result in large amounts of information being returned in the search results and a lot of time is spent browsing the outputs to identify relevant information that meets their needs. For example, a searcher who has expected to finish his search in 10 minutes unfortunately can take up to 30 minutes to obtain the desired results. The searcher may feel frustrated during the search process and may give up the search process before completing the task.

3. Terminology and Search Terms

According to Colomb (2002), we search for information when we know what we are searching for. On the other hand, we only browse for information when we do not really know what kind of information we really want. Keyword search is a common method used in information retrieval. However, keyword search does not take into account the meaning of words or semantic relationships between resources. For example, when a user wishes to search for information on the history of Myanmar, the search will only return resources that have keywords such as 'history' and 'Myanmar' and will not return information relating to the history of Burma although Myanmar is formerly known as Burma. In another example, consider a user who is searching for information on Kuala Lumpur. The results may return information that contains Kuala Lumpur as the capital of Malaysia, Kuala Lumpur as a destination for travel, Kuala Lumpur International Airport, or images of Kuala Lumpur. In this case, the search results match the search query; however human interpretation is still required to identify information that the user seeks which may be places of interest in Kuala Lumpur.

Yao (1995) states that the relevance of documents is essential to the effectiveness of information retrieval systems which often requires a user's judgement on whether the document is relevant to meet user's needs. In other words, human interpretation of searched results is required to determine the relevance and usefulness of the results. Consider an example that the phrase "Dora the Explorer" is entered in the search engine in the Web. The search results may return outputs relating to books, videos, toys, games, or merchandise of Dora the Explorer. Thus, in order to find the desired information (in this case, assume that we are interested in the characters of the television show Dora the Explorer), we have to browse the returned results to search for information on the characters of the show.

4. Synonyms and Homonyms

Henzinger (2007) states that synonym problems happen when users have an expectation that their query terms always appear in the search results. It is suggested that alternate query terms should be used as synonyms in addition to the first word used to allow more choices of results to return. For example, terms such as 'products' can also be searched using synonym terms such as 'artefacts' or 'goods'.

5. Lack of User Experience

Chowdhury (2004) explains that users may encounter problems when attempting to put together a search query using the combination of logical operators such as AND, OR, and NOT. If these operators are used incorrectly, it may result in outputs that are too narrow or too broad. Weinberg (1995) explains that a Boolean search can be used in free-text searches to relate two or more terms. However, these terms may not match with any words that the user hopes to find in the database. Thus, lack of user experience can result in inappropriate query formulation.

3.2 An Ontology-based Search Framework

Problems associated with information retrieval as described above can be overcome by using an ontology-based search approach. The use of ontology can help to formulate a query, which can be organised by defining a relationship of concepts using a hierarchical structure in the form of class/subclass relationship. Ontology for

a domain of specific content can be developed to be used to formulate a query. The ontology datastore can also be restructured easily when a relationship of concepts needs to be updated. With the proposed ontology approach, database content can be referred to by defining properties and relationships for instances/individuals in the datastore.

Figure 3.1 shows the processes when ontology-based search framework is applied. In this case, the user issues a query to the query interface. The query interface will perform the search by formulating a query based on the ontology datastore structure. The query interface is designed to match the hierarchical structure of the ontology datastore in the form of classes/subclasses relationship, which represent keywords/sub-keywords structures in the user query interface. After the query is executed, matched documents and results are retrieved and outputs will be sent to the user to review. If a desired result is not returned, the user repeats the query and the process continues until a desired result is obtained.

Figure 3.1: Ontology-based search framework

A travel domain (as shown in Figure 3.2) will be used to illustrate how the framework can be applied. In this example the concept of travel is categorised into five main classes: *accommodation*, *accommodation rating*, *activity*, *destination* and *contact*. Each category of class consists of more subcategories. For example, the *activity* category consists of four subcategories: *adventure*, *relaxation*, *sightseeing* and *sports*, while the *destination* category consists of eight subcategories: *backpackerDestination*, *beach*, *BudgetHotelDestination*, *FamilyDestination*, *QuietDestination*, *RetireeDestination*, *ruralArea* and *urbanArea*. This ontology will function as a dynamic query interface to assist users to formulate a query easily. This way, the problem of query formulation of not knowing the right keywords to use can be resolved. This is in contrast to open keyword searches in which the user needs to be skillful in knowing which keywords or Boolean operators to use in searching. The

hierarchical structure of ontology also presents a dynamic user interface that is easy to use to assist users and it can reduce time to formulate a query. Thus, the searcher needs not to be searching the database based on keywords that were included in the index of the database system, as keywords have been identified and defined using class/subclass hierarchical structure in ontology. The searcher can formulate his or her query based on the ontology of keywords that has been identified. This allows the searcher to search based on the searcher's own query formulation to determine how the search should be conducted. We call this the dynamic search, which offers flexibility on how a query is formulated and made.

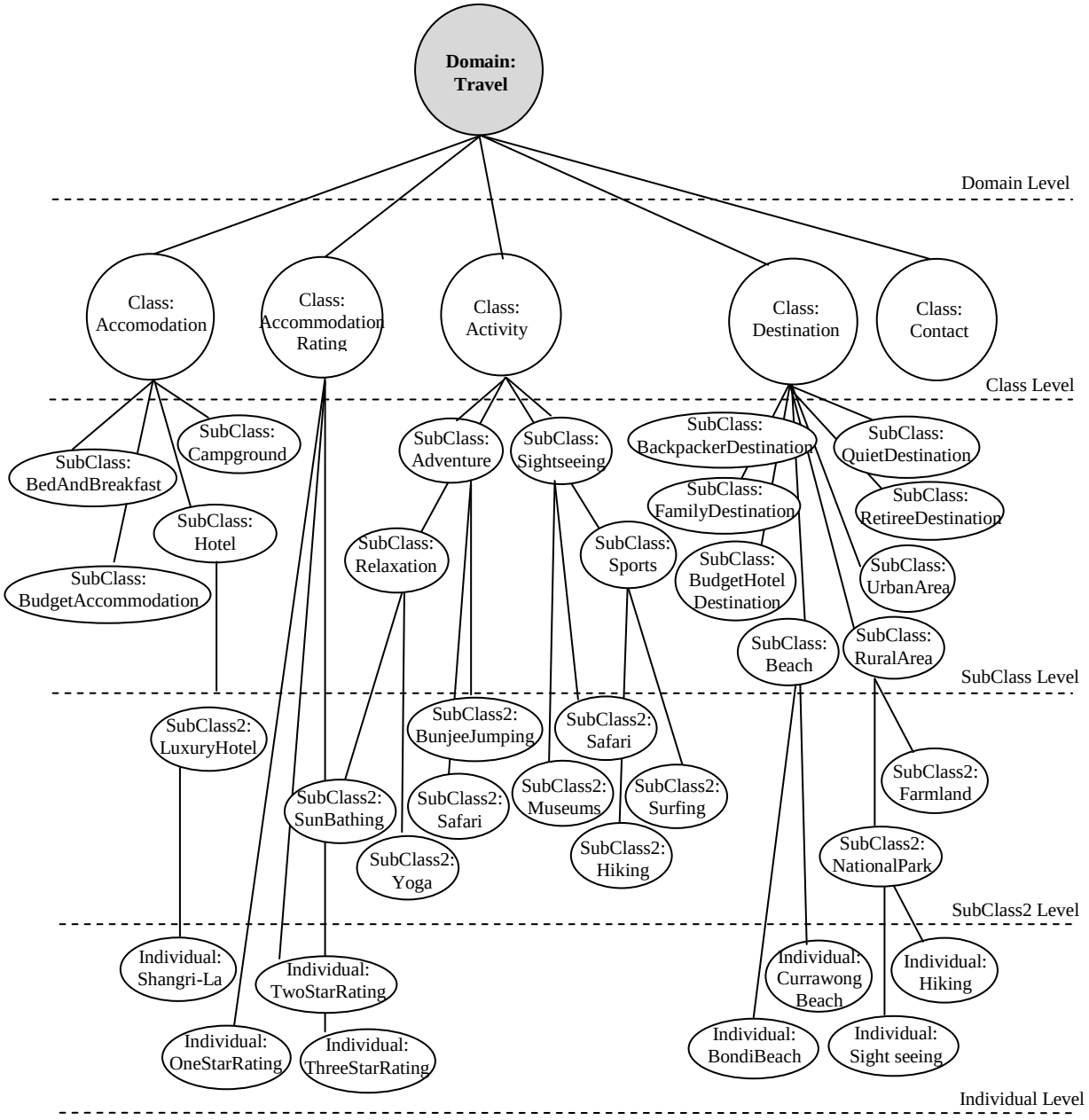


Figure 3.2: Ontology class-individual diagram

Another advantage of using ontology to formulate a query is to enable the searcher to have a unified structure when performing the search process based on an ontology class-individual diagram (Figure 3.2). For example, when the searcher wishes to search for accommodation and what he/she has in mind is *"I am interested in staying in budget accommodation with a 3-star rating".* This query can be formulated in the following hierarchical form:

Travel

Accommodation

BudgetAccommodation

Let us consider another example, where the searcher wants to search to compare "*interesting sports to do at the beach*". Using the above ontology the search can be formulated as:

Travel

Activity

Sports

This way the searcher does not need to use the Boolean operator as the search results are narrowed when the hierarchical structure of the query is executed. Using the ontology hierarchical structure, classes and subclasses can be constructed based on a mind-map (also known as a concept map). The mind-map is generated based on a human idea and a human's perception of the way information can be represented to show how information is related to each other. The elements of a mind-map are arranged according to major/broad concepts and are narrowed down to the minor/specific elements. In general, the major/broad concepts are placed on the top of the map followed by concepts that are more specific. Novak and Cañas (2006) explains that mind-maps or concept maps can be arranged hierarchically with lines or arcs connecting between two or more concepts to show the meaning of the relationship. Figure 3.3 shows the conversion of a mind-map to an ontology-based structure for the above travel domain.

In addition to the class/subclass hierarchy, properties can be used to describe relationships between classes and subclasses. There are two main types of properties: object properties and datatype properties. The object properties link individual instances. Consider "is-a" relation property in which the class/subclass hierarchy is represented by "is-a" relation. A class "A" is-a subclass of "B" if every instance of "B" is also an instance of "A". For example, "*hotel"* is a subclass of

"accommodation". This way searching for hotel information can be performed by searching for *accommodation* class first and then narrowing down to *hotel* subclass.

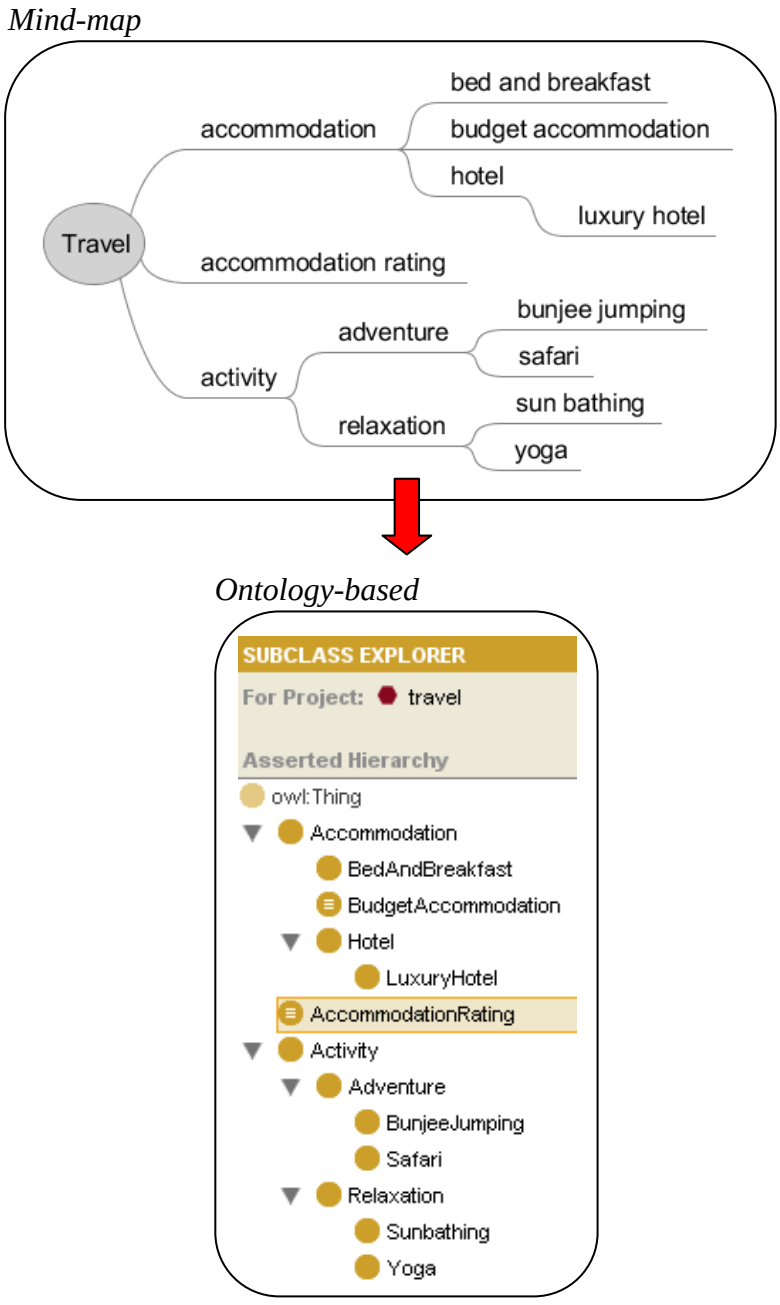


Figure 3.3: Mind-map to ontology-based structure

Using ontology, synonym words can be established using the functional properties featured in ontology. For example, a "*5-star rating hotel"* and a *"Luxury hotel*" refer to the same type of hotel. In this case, we use a single value property to show that a

5-star rating hotel is a synonym for *luxury hotel*. Figure 3.4 shows the example of a synonym composition.

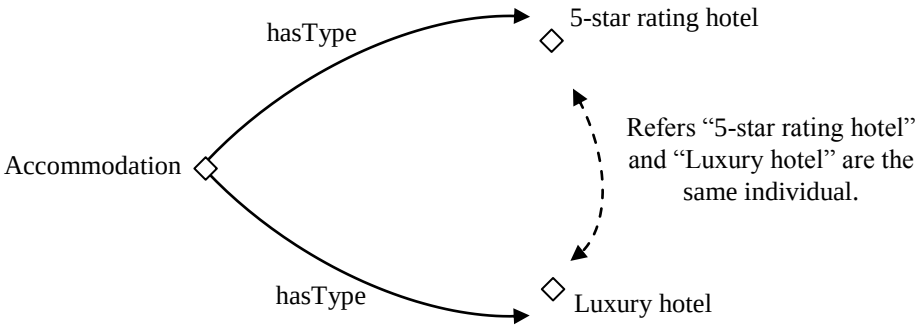


Figure 3.4: An example of a synonym composition

Relationships of the data ontology can be defined by using the data properties features. This feature describes relationships between individuals and data values (see Figure 3.5). For example, we use *hasRating* to show relationships between the class *accommodation* and the class *accommodationRating*.

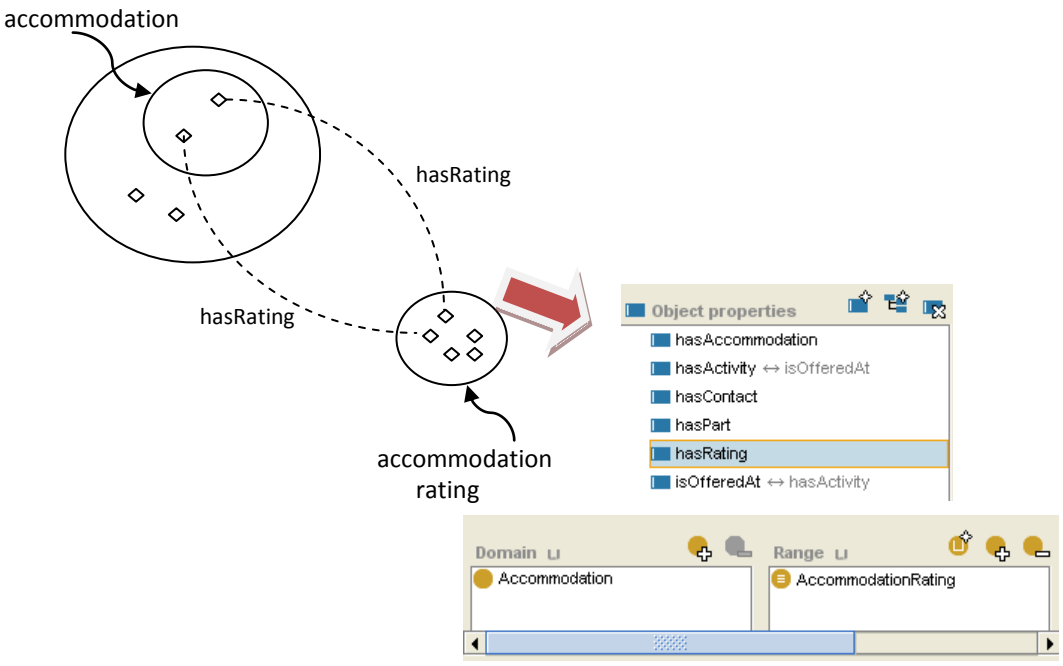


Figure 3.5: An example of relationships classification

After the information is presented in the ontological hierarchical structure, the ontology user interface can be designed to present as dynamic data interface in the form of drop-down menu features. The main advantage of using a drop-down menu is that the searcher does not need to think of keywords to use; this helps users in a query formulation.

For example, Figure 3.6 shows a drop-down menu data structure that corresponds to the ontology datastore of the example travel domain. Here it can be seen that the construction of the drop-down menu interface follows the hierarchical structure in ontology datastore.

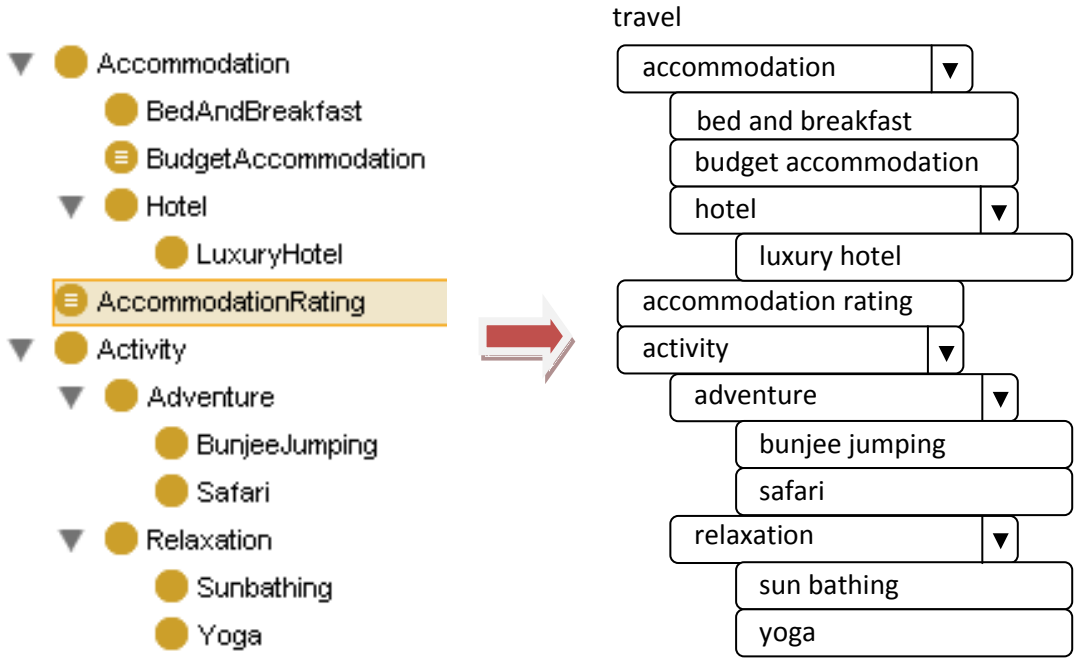


Figure 3.6: Ontology-based to drop-down features

3.3 Conclusion

In this chapter, major retrieval issues on conventional online database search systems faced by novice researchers are described. An ontology-based search framework is proposed and discussed. The framework consists of five elements: (1) user, (2) query, (3) query interface, (4) search process, and (5) result. The user conducts query

formulation through the query interface. The query interface is designed based on the ontology hierarchical structure in the ontology datastore and is represented as keywords/sub-keywords structures in the query interface. The system will process the search query once the query is sent. The search process will look for the matched documents and return results to the user. The results that are returned to the users will be reviewed and users have the option to repeat the process if the results do not fulfil the query made.

The application of the framework on travel domain has been used to explain how the dynamic search environment is developed to help in query formulation. Ontology enables relationships between keywords and terms to be defined. Ontology allows desired information to be retrieved by sharing common vocabularies with an understanding of meaning of terms in the domain.

In the following chapter, a prototype system is developed based on the design science methodology selected to guide the system development. The system development takes into account problems identification in the database search environment and motivation in improving and extending current limitations. An illustration of the demonstration of the applied approach is described, followed by the evaluation of the prototype system.

CHAPTER 4: ANALYSIS OF THE ONTOLOGY-BASED SEARCH FRAMEWORK

This chapter presents the application of the ontology-based search framework to a case study. A prototype system is developed by using the design science method proposed by Ken *et al.* (2007): (1) problem identification and motivation, (2) objectives of solution, (3) design and development, (4) demonstration, (5) evaluation, and (6) communication.

The content of this chapter is organised as follows: Section 4.1 discusses the design science methodology in information system research. In this research, the six steps approach proposed by Ken *et al.* (2007) is adopted. Sections 4.2 to 4.7 explain the six steps of the Ken *et al.* (2007) approach. Section 4.2 discusses problem identification and motivation. Section 4.3 lists the objectives of solution to overcome the problems. Section 4.4 describes the design and development process of the ontology-based search system. Section 4.5 illustrates the demonstration of the applied approach. Section 4.6 presents the evaluation process and reports the outcome of data analysis. Section 4.7 lists the communication results. Finally, Section 4.8 concludes this chapter.

4.1 Design Science Methodology

Research design becomes a practice to all researchers in developing research information for a particular design project (Sato, 2004). In information system (IS) research, behavioural science and design science are the two significant paradigms. Behavioural science is focussed on theories based on human or organisational behaviour. Hevner *et al.* (2004) differentiate between behavioural science and design science, as behavioural science research paradigm is concerning "What is true?" while design science should answer the question of "What is effective?" Hevner *et al.* (2004) define the behavioural science as a paradigm that explains and predicts human or organisational behaviour through the define theories. March and Smith (1995) define design science as the artefacts or technologies which are developed to

fulfil human purposes. In design science research, there are two activities identified as basic activities: (1) build and (2) evaluate. The building process is to determine artefact construction purpose and its development, while the evaluation process is to verify the artefact performance.

Birkhofer (2011) expresses design science as a part of theories, paradigms, models, methods and knowledge. Hevner *et al.* (2004) highlight that the importance of design science paradigm is to produce an artefact to address a problem by taking in the knowledge and understanding of a problem domain. According to Hevner *et al.* (2004) the problem solving process includes seven guidelines: (1) design as an artefact, (2) problem relevance, (3) design evaluation, (4) research contribution, (5) research rigor, (6) design as a research process, and (7) communication of research.

The first guideline is described as a main point in the design and implementation process of the appropriate output by addressing and analysing the problem to be solved. Guideline 2 is aimed to apply the technology-based solutions to the problem identified. In guideline 3, design evaluation will determine the possibility of the design's usefulness, efficiency and whether it's worthy to be completed through an evaluation process, which provides the necessary feedback to help in the development process. Guideline 4 is identifying the contribution of the artefacts in the area of design construction knowledge and/or design evaluation knowledge. However, there are two criteria which the artefact's contribution is being assessed to ensure the proposed design accurately represents the case environments and must be implementable. Guideline 5 is which variety of method in the design artefact's construction and evaluation is needed. However, the important part of running the evaluation process is to determine how well the artefact is working. Guideline 6 is the step in analysing the best design process that can fulfil the desired end project. Lastly, guideline 7 is the final process to apply the design artefact in the real organisation with the sufficient information provided.

In this research, we adopt the six steps of design science research process introduced by Ken *et al.* (2007) when applying the ontology-based search framework to the case study. These six steps are: (1) problem identification and motivation, (2) objectives

of solution, (3) design and development, (4) demonstration, (5) evaluation, and (6) communication as shown in Figure 4.1. Ken *et al.* (2007) also include three elements in the design science research methodology's characteristics: (1) principles, (2) practices, and (3) procedures. These elements are important to succeed in the creation and presentation of design science research artefacts. Principles of the design science research refer to inclusion of any designed object to understand research problems; practices describe the problem solving process; and procedures provide an acceptance process to conduct research. Ken *et al.* (2007) further include two models: (1) nominal process and (2) mental models. The nominal process model provides guidance on the design science research output while the mental model helps in recognising and evaluating results of research output. By following this model, the researcher can understand and evaluate the work of others to determine the research potentiality. For example, in the research process, the researcher can refer to other research outcomes or previous research outputs to help in expanding current research limitation. Without the nominal process and mental model, researchers have difficulties to distinguish from prior research.

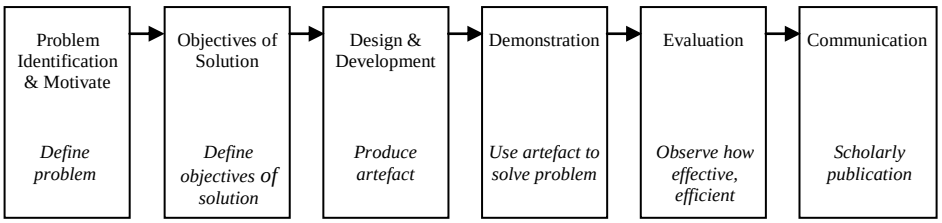


Figure 4.1: Design Science Research Methodology

4.2 Problem Identification and Motivation

The case study concerns problems faced by novice researchers such as university students. In this research case, university students who embark on a research programme for the first time, often find identifying a viable research topic to be a daunting process. Choosing a research topic for a thesis and dissertation is a complex process. The students not only need to have some interests in the topic area, they also need to choose a topic area that will make a contribution to the area of research. It is

not uncommon for students to seek advice from academic staff from the department they are enrolled in; they will also search existing literature to determine how much information and research is already available on the topic. Usually students also search previous dissertations or theses available in the library to identify a potential research topic. This approach not only functions as a good starting point, it also enables students to develop an initial focus on the research topic that they wish to embark on. This is particularly important for students who have to complete the thesis in the final year of their course such as the student in the Honour programme who only has one year to complete the thesis. Very often these students are novice inexperienced researchers who may not be skilful in academic search. Although students have undergone at least three years of undergraduate study before embarking on the research programme and have acquired some degree of literature search skills or using library services, very often their literature search experience is limited. For example, students are often given a list of bibliographies or reference lists in the subject area by their lecturers, which provide a good starting point when a literature search is conducted. This may not be the case when students need to search for a thesis topic that is of interest to them as well as a doable research topic that can be completed within the specific timeframe of their degree. Thus, a search and retrieval system that can help students in this initial phase of identifying the research topic is desired.

4.3 Objective of Solution

As explained in the statement of problem section in Chapter 1, students face difficulty in formulating the right keywords in a search query; search results often not matching with the expected search outcome; search results do not match with desired topic area; having difficulty in cross-disciplinary topics and searches can only be performed in a mono-language query. These problems can be categorised into five groups of problems: (1) query formulation, (2) lack of user perseverance, (3) terminology and search terms, (4) lack of user experience, and (5) synonyms and homonyms.

In this research, ontology-based search framework has been proposed in Chapter 3. The proposed framework will be applied to overcome the identified problems. The objective of the proposed system is to develop a search system to help students in query formulation to search thesis reports (Honour, Master and PhD) for the Education Faculty at Universiti Teknologi Malaysia (UTM). The thesis resource aims to assist students with a quick and easy reference point in which students can search for past completed theses. One of the aims of the system is to assist students who are novice researchers in identifying research topics that have been completed in the past years. The semantic search framework is adopted to overcome problems encountered in keyword searches such as synonym, short-form term and bilingual query.

4.4 Design and Development

The ontology-based search system is developed based on an ontology-based mind-map. The mind-map is developed from the academic programme profile of the faculty in this case study. The Education Faculty aims to produce future teachers with knowledge and experiences related to the teaching profession. Thus, research topics are often conducted on issues related to teaching and learning based on specialised and professional subjects offered at the faculty. The mind-map is developed by considering the relationship between major components of teaching and learning as specialised subject offering, for example mathematics, physics, chemistry, living skills, sports science, Islamic studies, computer science and Teaching Language as a Second Language (TESL).

The mind-map is organised in a hierarchical structure to be translated to an ontology structure. Figure 4.2 shows a sample snapshot of a mind-map expressed in the English language, and Figure 4.3 shows a sample snapshot of the same mind-map in the native language (Bahasa Malaysia). Appendix A shows the complete table of ontology data. The mind-map is developed using the inductive approach of ontology design (Holsapple and Joshi, 2002). With an inductive approach, the researcher observes, examines and analyses the sample domain of interest to develop the required ontology.

To prepare for mind-map development, the author examines the research and teaching areas in the faculty. Five main categories of teaching and learning that match the faculty profile are identified: teaching, learning, field of study, education level and others. For each main category, subcategories are prepared to show relevant topic areas in each category. For example, the *teaching* main category consists of the following subcategories: *pedagogy, educator, skill, style, course or subject, theory* and *tool*. To illustrate the subcategory of *pedagogy* relates to research conducted on a thesis, which investigates how the teaching pedagogy is applied. The same rationale applies for the *learning* category. In the case of *field of study* category, it reflects the course or subject offered in the faculty. These include *mathematics, physics, chemistry, living skills, sports science, Islamic studies, computer science* and *TESL*. The same category is also included as subcategories of *teaching* and *learning*. These examples of cross-categories options enable users to select the same keyword, but under a different structure of category-subcategory, to give a variety of combinations to enable users to choose from different perspectives.

As explained, the proposed ontology development is based on inductive approach which is based on the observation and analysing of current or specific domains to apply to particular domains. The prototype development is based on existing library records in the database system of Faculty of Education, Universiti Technologi Malaysia The thesis records obtained for the prototype development was derived from the database system in 2008. Thus using the inductive approach the subcategories created in the prototype system are based on existing structure of the database system. It is worth noting that the subcategories structure can be changed whenever a new structure of categories and subcategories is needed or requires modification.

Another unique characteristic found in this university is that the thesis can be written in the native language of Bahasa Malaysia or the English language. As a non-English speaking country, this university conducts most of the courses in the Bahasa Malaysia language. However, courses can be taught in the English language, particularly for students who are enrolled in English teaching courses such as TESL or international student. These students write their thesis in English; therefore it is

desirable for the search system to be able to conduct searches for thesis records regardless of language.

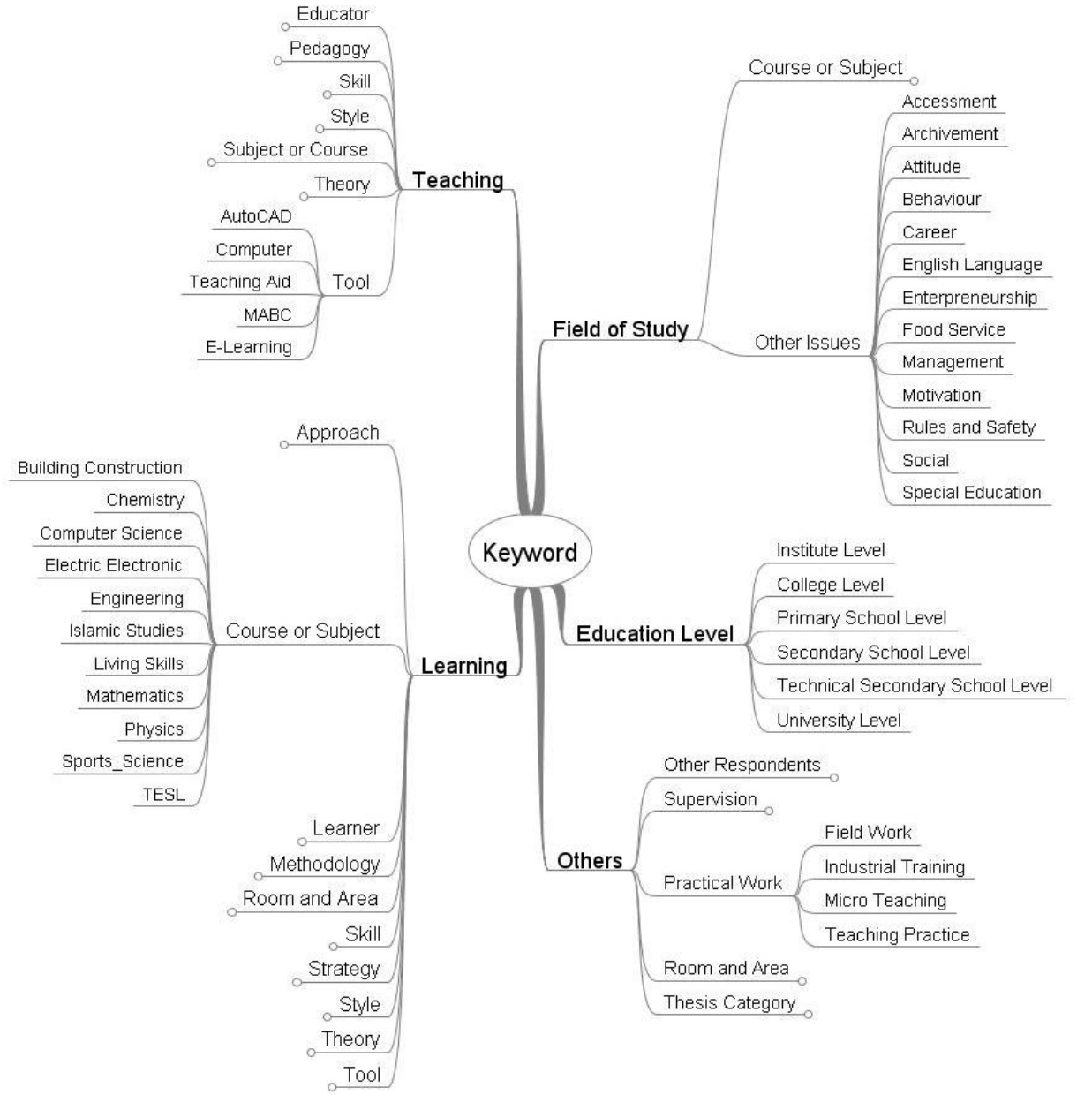


Figure 4.2: A sample mind-map of thesis ontology in the English Language

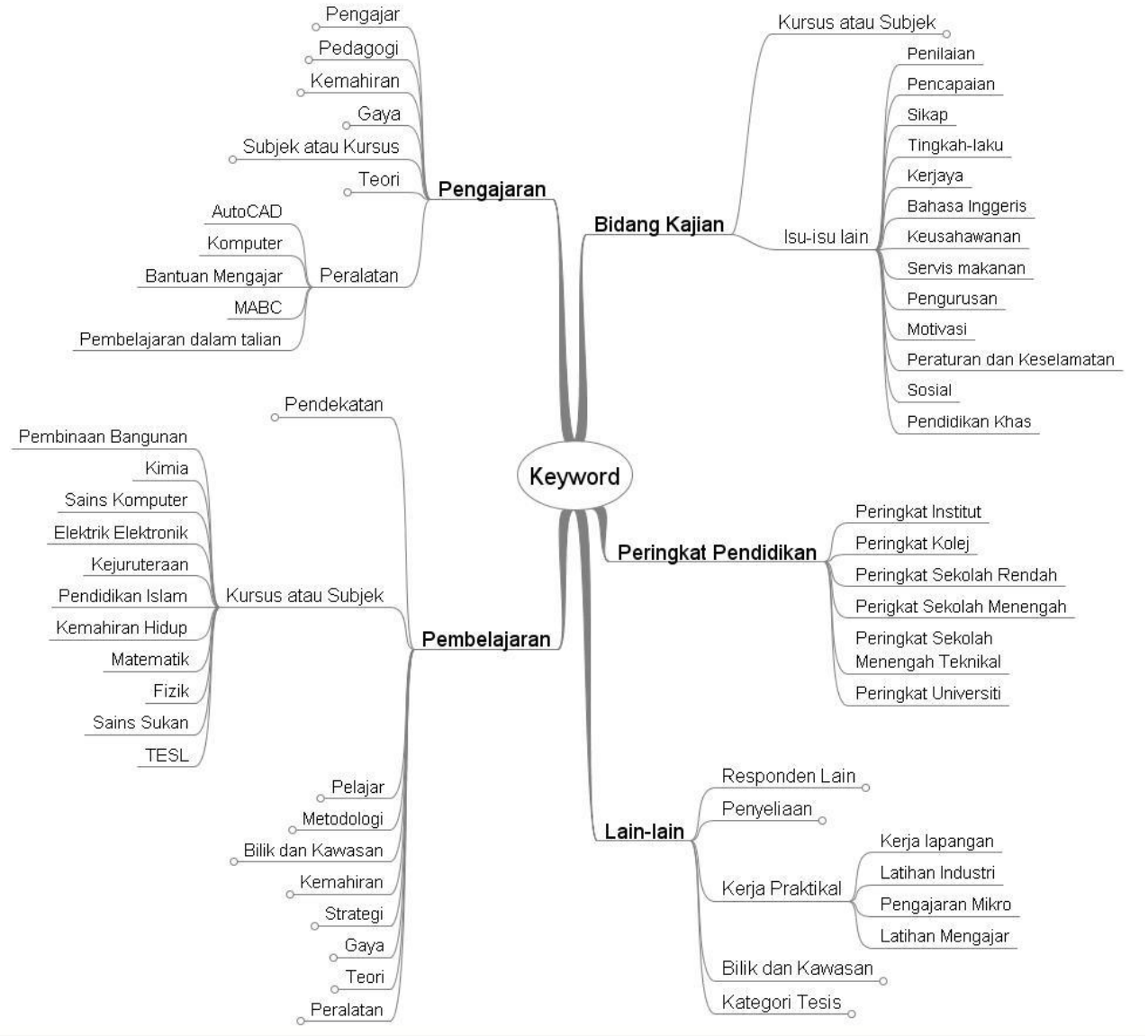


Figure 4.3: A sample mind-map of thesis ontology in the Bahasa Malaysia Language

Protégé 3.4.4 is used to develop the ontology classes and subclasses. Protégé is a tool to help users in the construction of a small to large knowledge repository. Protégé represents the concepts and relationships as mind-map or concept map features, using an easy to use graphical representation and it allows users to have direct manipulation such as content controlling and editing. This way, the knowledge browsing and editing process can become simple and flexible. Figure 4.4 shows a partial snapshot of ontology classes/subclasses corresponding to the mind-map as shown in Figure 4.3.

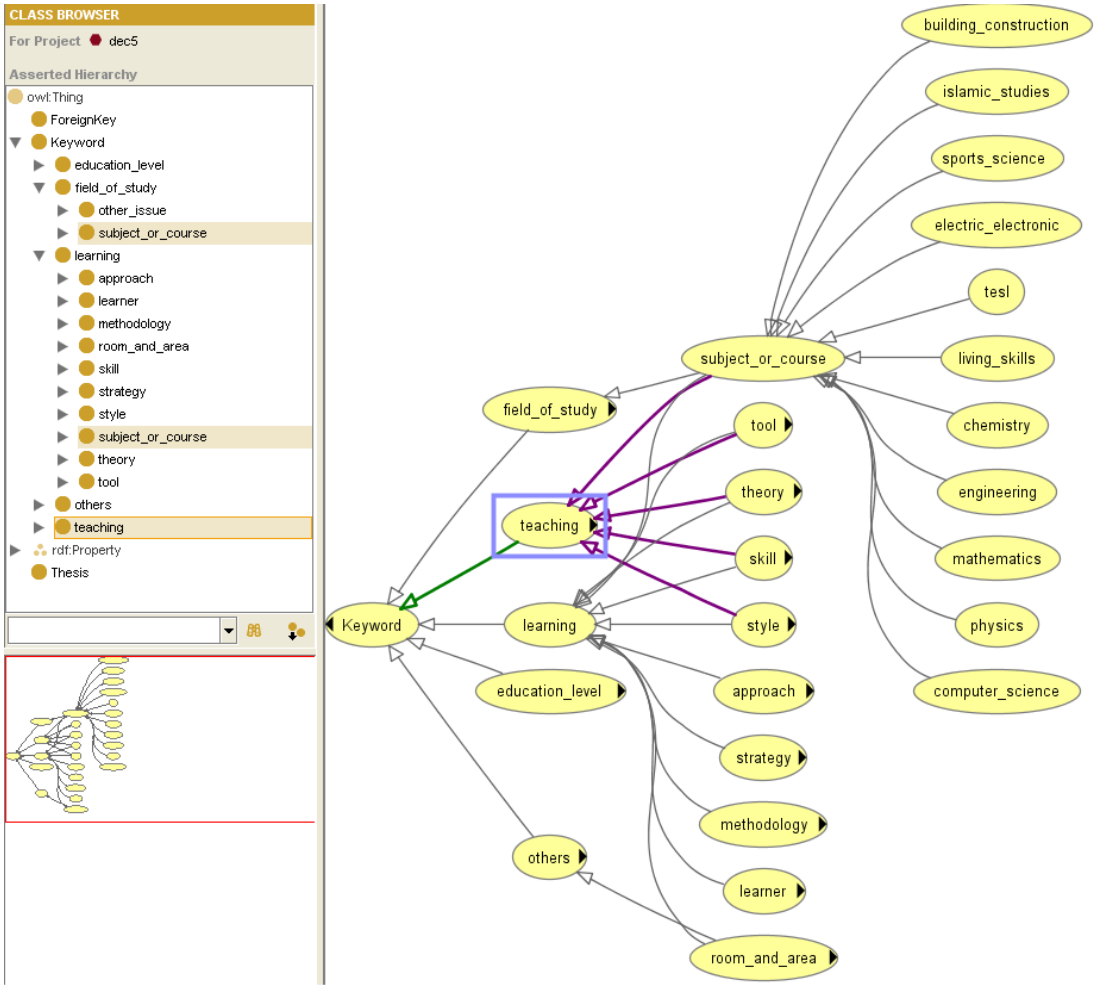


Figure 4.4: A partial snapshot of hierarchy of thesis ontology

To illustrate the application of the ontology-based search framework, consider an example when a student conducts a query, *"I am interested in investigating issues on the use of AutoCAD in teaching and learning at school"*. This query can be expressed in the following hierarchical form based on the mind-map developed:

Learning

Tool

AutoCAD

However, other students may see this query in different perspectives where they may choose a different query in the hierarchical form such as:

Teaching

Tool

AutoCAD

This way the students are free to select keywords that correspond to their thinking or mind-map. Thus, the ontology-based approach offers flexibility on how a query is made and it more closely resembles how a human thinks when they formulate their search query.

Consider another example, when the user is searching for past research topics that are related to *"teachers' perceptions of delivering lessons using computer assisted learning"*. Using the ontology as described in Figure 4.4, a search query can be formulated as follows:

Teaching

Tool

Computer

Consider another example, where the user wants to search *"I wish to compare learning styles of second year education students with the teaching styles of their lecturer".* Using the ontology the search query can be formulated as:

Teaching

Style

Or

Learning

Style

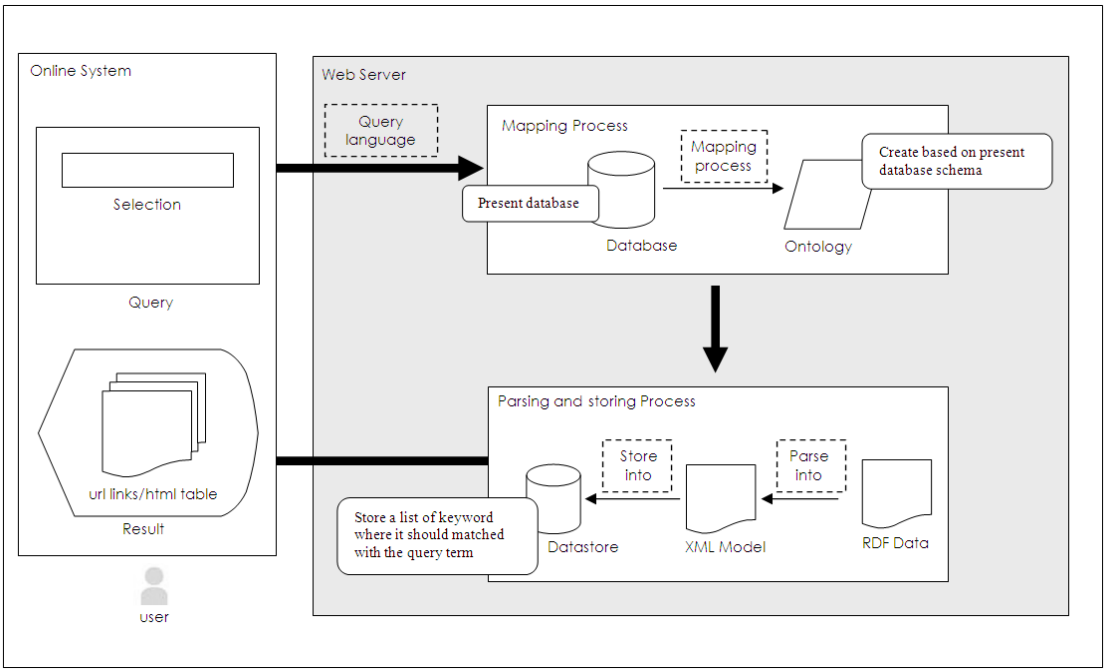


Figure 4.5: System development design

Figure 4.5 shows the system development design of the proposed system. Users can access the system through a Web server where the appropriate user interface is created. When the ontology datastore is complete; it will be converted into a Resource Description Framework (RDF) document, which is then parsed and all data is dumped into a relational database. In this system, Apache Tomcat is used as a Web services.

The thesis records which are stored in MySQL database are being re-used where the process of mapping the database records into ontology class/subclasses is required to transform the existing database record into an ontology database. This process of importing data from MySQL into Protégé is carried out using Protégé Plug-in called DataMaster (Nyulas and Tu, 2007). The DataMaster plug-in can be used with any relational database with JDBC/ODBC drivers to import schema structure and data from relational databases into Protégé. The importing process is done through its user interface where a connection to relational database has to be login first. In this prototype, only some of the table content is imported to map into the pre-defined ontology classes and subclasses.

The next stage of document parsing is using the PHP code to store RDF documents into a relational database. RDF documents are needed to be parsed and stored as triple store in any relational database. This process allows ontology sharing and queries to be made more easily by using a standard MySQL query. The retrieval process of the RDF documents once it has been parsed into MySQL database is organised into three columns of data, which are subject, predicate and object. The process of parsing is needed once only as long as there is no change in the class/subclass hierarchy and no new data/records are added, the query can then be run.

4.5 Demonstration

The main features of the system are to allow shortened forms of terms, synonyms and bilingual searches. We use the following three scenarios to show how the system works.

4.5.1 Scenario 1 - Information Retrieval based on Shortened Forms of Terms

In this scenario a student likes to find any research projects that are related to the subject “English for Academic Communication”. The query result can be retrieved with the keyword ‘UHB1412’ (which is the course code) or subject name “English for Academic Communication”.

Figure 4.6 shows how the relationship can be represented between these two words which we have defined using the synonyms relation of ‘SameAs’.

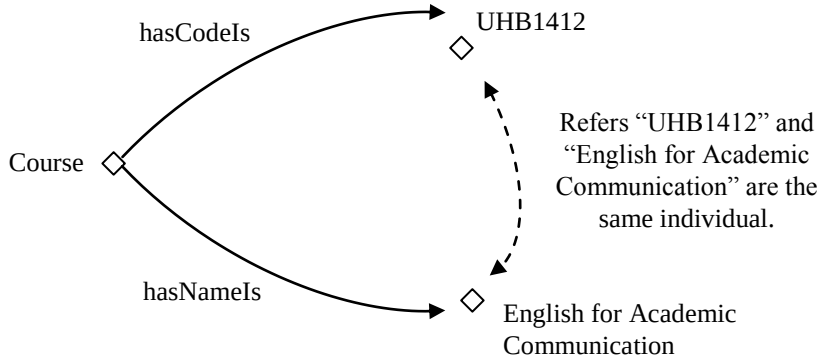


Figure 4.6: Relationship of "SameAs" Individual: Shortened Forms of Terms

4.5.2 Scenario 2 - Information Retrieval based on Different Languages

In this scenario, the student wishes to identify all theses, written in both Bahasa Malaysia and English language, for research topic on 'teaching' category. For the local students they will choose to use the keyword 'pengajaran' (meaning teaching in Bahasa Malaysia language), whereas for international students they will use the keyword 'teaching'. Thus, the query that can handle both languages desired. Figure 4.7 shows how the relationship is assembled.

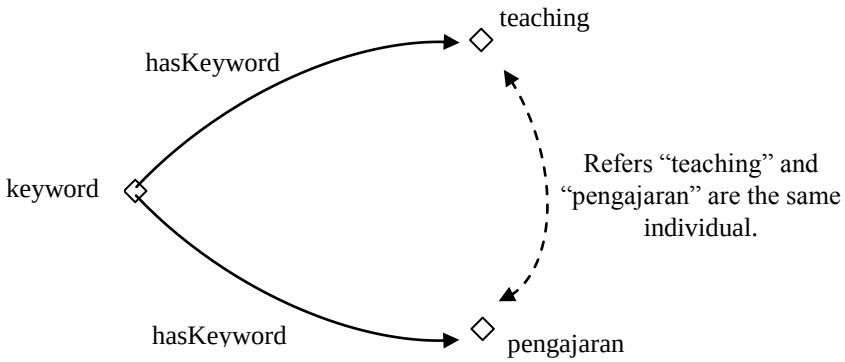


Figure 4.7: Relationship of "SameAs" Individual: Bilingual

4.5.3 Scenario 3 - Information Retrieval based on Synonym Relation

In this scenario, the student likes to conduct a project development of smart applications and would like to know whether any previous research has been conducted in this area. Thus, the student will think of possible keywords such as 'smart application'. However, the keyword of 'smart' is not the only word that can

describe smart applications. There are some other words that are capable of describing smart applications such as AI (artificial intelligence) or intelligent. In this case, we use a synonym relation to give more meaning to the data.

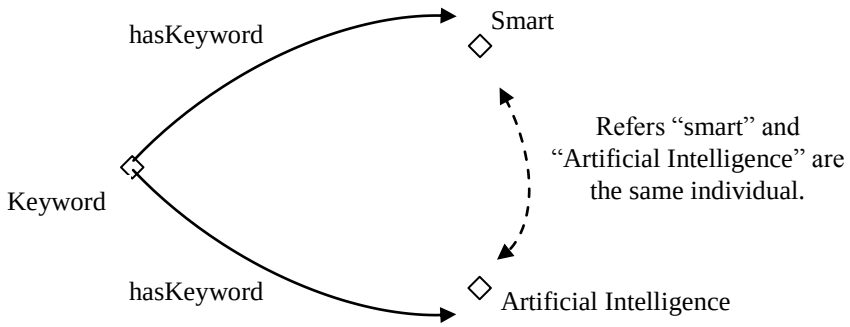


Figure 4.8: Relationship of “SameAs” Individual: Synonym

Figure 4.8 shows how the relationship can be represented between these two terms which we have defined using the synonyms relation of ‘SameAs’.

4.6 Evaluation

Design evaluation is very important to observe how effective the system is. The artefact needs a kind of measurement tool and has to be analysed to establish whether the artefact can support solutions to the problem identified previously. Different types of measurement tools can be used to achieve a variety of feedback, such as having a short interview with respondents after using the artefact; recording activities undertaken while the artefact is being used by the respondents (such as a log book); or distribute questionnaires to evaluate usability, practicability and compatibility of the proposed artefact to the target environment.

In this research case, evaluation by questionnaire is conducted. The evaluation process is described as follows:

4.6.1 Evaluation based on Student Survey

The purpose of the evaluation process is to enable current students enrolled in the Education Faculty, Universiti Teknologi Malaysia, to evaluate the prototype system. The respondents are Year 1 students who have enrolled in the degree programmes at the Faculty of Education, Universiti Teknologi Malaysia. These respondents are novice researchers who have only basic research skills and are inexperienced in academic research. There are 191 Year 1 students enrolled in Semester 1 in the 2011/2012 academic years. Only 133 (69.6%) students attended the laboratory classes on the day of the survey and only 110 (82.7%) usable responses have been included for analysis; 23 (17.3%) responses were rejected due to incomplete answers.

In the scenario-based questions, only 110 usable responses have been included in the analysis of result. Although each student is given 3 attempts for each scenario, however majority of students have only attempted the search once for each scenario. For this reason the number of attempts shown in Table 4.7 to Table 4.9 did not show a total of 330 attempts. For example only 198 search attempts have been conducted for Scenario 1. This shortcoming can be overcome in future research by providing more explicit instruction to students participated in the survey to ensure all students performed 3 search attempts in the evaluation of the prototype system.

Due to the bilingual nature of the students enrolled in the Faculty, the questionnaires were distributed in two languages: English and Bahasa Malaysia. However, respondents have chosen to complete the questionnaire in the Bahasa Malaysia language. The evaluation of the prototype was conducted in the computer laboratory. The survey was supervised by a lecturer and one technical support staff member was assigned in each laboratory (three computer laboratories were used during the survey).

To complete the survey students are asked to access to the server using the link http://web2.fp.utm.my/syikin/index.php to use the prototype. The server is provided by the Faculty of Education. The system has been uploaded remotely using an ftp server which is WinSCP version 4.3.5. The prototype system has been uploaded to

the server a month before the survey to allow the system to be tested and be functional before the student evaluation is conducted.

The survey contains three sections; Part A, Part B and Part C. Part A consists of questions to identify the background of the respondents. Part B presents three scenarios in which the students were asked to perform search queries using the prototype system. Students are asked to repeat the search query for each scenario with up to three attempts. Depending on the selection made, students may or may not find the expected query results. Students are asked to indicate whether the expected output is obtained.

In Part C, students have to answer 18 Likert scale questionnaires. The Likert scale is made up of 1 to 5 scales. 1 refers to "strongly disagree", 2 refers to "disagree", 3 refers to "slightly agree", 4 refers to "agree" and 5 refers to "strongly agree". These questions aim to evaluate the student's experience in using the prototype system. The questions relate to the ability of the system to facilitate the search process and the user friendliness of the system. It is worth noting that students are asked to answer Part A before answering Part B then followed by Part C.

There are negative and positive statements used in the questionnaire. Adams and Cox (2008) explain that it is important to include both negative and positive statements in the questionnaire to determine whether respondents were careful in reading and answering the questions. This is because the order of questions presented may result in bias; respondents may give constant responses (for example, respondents may tend to select most of the answers in positive responses, negative or neutral responses). To counter this problem the questions in the questionnaire should be stated in the form of negative and positive statements. The questionnaire is shown in Appendix C.

4.6.2 Results of the Survey

Tables 4.1 to 4.6 show results related to background of respondents. In the survey 64.5% of respondents are female and the majority of respondents (61.8%) are between 18 to 20 years old. Based on student enrolments, 40% of respondents possess a Diploma education level and 37.3% of respondents have passed their

Matriculation level. In term of the computer skill background, 60.9% respondents are categorised as “skilled” and 83.6% of respondents have more than 3 years experience using the Internet. For Internet usage frequency, 53.6% of the respondents use the Internet everyday.

Table 4.1: Respondent’s Background: Sex

Sex	Number of respondents	Percentage
Male	39	35.5%
Female	71	64.5%
Total	110	100.00%

Table 4.2: Respondent’s Background: Age

Age	Number of respondents	Percentage
Under 18	0	00.0%
18-20	68	61.8%
21-23	37	33.6%
24 and above	5	4.6%
Total	110	100.00%

Table 4.3: Respondent’s Background: Current education level

Current education level	Number of respondents	Percentage
After SPM	1	0.9%
After STPM	24	21.8%
After Matriculation	41	37.3%
Diploma	44	40.0%
Total	110	100.00%

Table 4.4: Respondent's Background: Computer skills

Computer skill	Number of respondents	Percentage
No skills	1	0.9%
Less skilled	0	0.0%
Slightly skilled	37	33.6%
Skilled	67	60.9%
Highly skilled	5	4.6%
Total	110	100.00%

Table 4.5: Respondent's Background: Internet usage

Internet usage	Number of respondents	Percentage
0 - 3 months	1	0.9%
3 - 6 months	3	2.7%
6 months - 1 year	3	2.7%
1 year - 3 years	11	10.0%
More than 3 years	92	83.6%
Total	110	100.00% *

*The total percentage is rounded to 100% as the total computed is 99.9%.

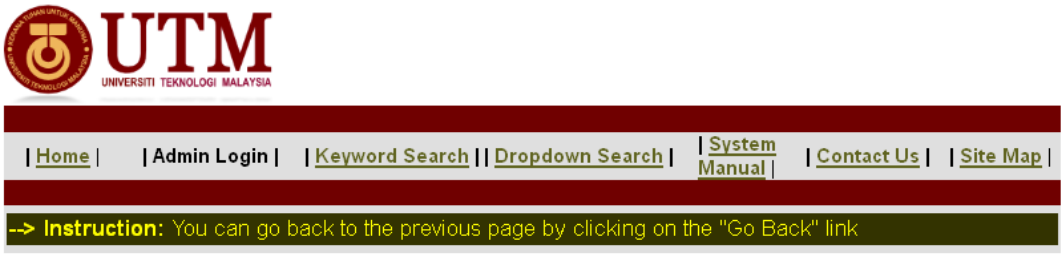
Table 4.6 Respondents Background: Internet usage frequency

Internet usage frequency	Number of respondents	Percentage
Sometimes in a year	0	0.0%
Sometimes in a month	2	1.8%
Sometimes in a week	21	19.1%
Everyday	59	53.6%
More than 1 time in a day	28	25.5%
Total	110	100.00%

Tables 4.7 to 4.9 show the results of keyword lists used in Scenario 1 to Scenario 3. For Scenario 1, respondents are required to search for a thesis containing the following thesis title:

> *"Lecturers' and TESL teacher trainees' perceptions towards the level of readiness in performing teaching practice".*

Figure 4.9 shows the expected result when the search is performed using the combination of keywords: "others → practical work → teaching practice". This combination is selected because the search related to teacher trainee and teaching practice. Although the keyword of "level of readiness" can be used to narrow the search it may result in too small a number of results returned. Therefore, the *"teaching practice"* keyword is selected as it is a major keyword that can broaden the results returned.



This search return **5 Rows** of results that matched with the query

« Go back

thesisNo	titleName
8017	PERSEPSI GURU PELATIH TERHADAP PENYELIAAN LATIHAN MENGAJAR OLEH GURU PEMBIMBING DAN PENYELIA UTM SEMASA MENJALANI LATIHAN MENGAJAR.
8019	LECTURERS' AND TESL TEACHER TRAINEES' PERCEPTIONS TOWARDS THE LEVEL OF READINESS IN PERFORMING TEACHING PRACTICE.
8040	PERSEPSI PELAJAR ALIRAN SAINS TAHUN DUA FAKULTI PENDIDIKAN DARI SEGI SIKAP, MINAT DAN PERSEKITARAN PEMBELAJARAN YANG POSITIF SEBELUM MENJALANI LATIHAN MENGAJAR
8047	STRES DALAM KALANGAN GURU PELATIH FAKULTI PENDIDIKAN UNIVERSITI TEKNOLOGI MALAYSIA SEMASA MENJALANI LATIHAN MENGAJAR.
8061	BIMBINGAN OLEH GURU PEMBIMBING KEPADA GURU PELATIH UTM.

Figure 4.9: Screenshot of expected result for Scenario 1

Table 4.7 shows the choices of combination of keywords used by the students in searching for the thesis title. The results show that 60.10% of respondents select the keyword "*teaching"* as their first choice of term in Scenario 1 followed by the combination of "*teaching → educator"* (33.33%) and "*teaching → educator → lecturer"* (11.11%).

Table 4.7: List of hierarchy of keywords used in Scenario 1

Row Labels	Data Number of Search Attempts	Percentage
education level	**11**	**5.56%**
field of study	**35**	**17.68%**
learning	**27**	**13.64%**
others	**6**	**3.03%**
teaching	**119**	**60.10%**
-	**40**	**20.20%**
trainee	1	0.51%
-	39	19.70%
educator	**66**	**33.33%**
lecturer	22	11.11%
mentor	1	0.51%
teacher	3	1.52%
teacher trainee	14	7.07%
technical teacher	5	2.53%
-	21	10.61%
pedagogy	**13**	**6.57%**
self teaching	12	6.06%
-	1	0.51%
Grand Total	**198**	**100.00%**

note: "-" sign means no option is selected

For Scenario 2, students are given the following scenario:

> *"You are required to do research for your course presentation. You have been assigned a topic related to 'teaching and learning in Mathematics'. You have to use the database system provided to search for those related theses and present the outcome about the choice of topics that have been done by the previous students".*

Figure 4.10 shows the expected result when searching using the combination of keywords as follows: *"field of study → subject or course → mathematics"*. This combination is selected as the subject name is known and is listed under *"subject or course"* keyword. However, if the searcher selects the first level keyword of *"teaching"* or *"learning"* this does not meet the requirement of the question, which should relate to both *"teaching"* and *"learning"* keywords. Thus, the first level of *"field of study"* keyword is a more general keyword to use in order to include teaching and learning topics.

Figure 4.10: Screenshot of expected result for Scenario 2

Table 4.8 shows choices of a combination of keywords selected by the students. The result shows that 53.70% of respondents select the keyword "*field of study*" as their first choice of term followed by "*field of study → subject or course*" (48.15%) and "*field of study → subject or course → mathematics*" (11.11%).

Table 4.8: List of hierarchy of keywords used in Scenario 2

Row Labels	Data Number of Search Attempts	Percentage
education level	**12**	**11.11%**
field of study	**58**	**53.70%**
-	**4**	**3.70%**
mathematics	1	0.93%
-	3	2.78%
other issue	**2**	**1.85%**
e-learning	1	0.93%
-	1	0.93%
subject or course	**52**	**48.15%**
building construction	1	0.93%
engineering	1	0.93%
mathematics	44	40.74%
-	6	5.56%
learning	**17**	**15.74%**
others	**4**	**3.70%**
teaching	**17**	**15.74%**
Grand Total	**108**	**100.00%**

note: "-" sign means no option is selected

For Scenario 3, students are required to conduct a search to find the following research's title:

> *"An investigation on approaches used to teach literature in the ESL classroom: A case study of Sekolah Menengah Kebangsaan Taman Desa Skudai, Johor Bahru".*

Figure 4.11 shows the expected result when searching using the combination of keywords as follows: "*field of study → subject or course → tesl"*. This combination of keyword is selected as teaching approach is identified as the main theme in this topic. However, if we choose to use the *"teaching → approach"* keyword

combination, the results will be too broad. By using the keyword combination of *"field of study → subject or course → tesl"* it will provide a more specific output.

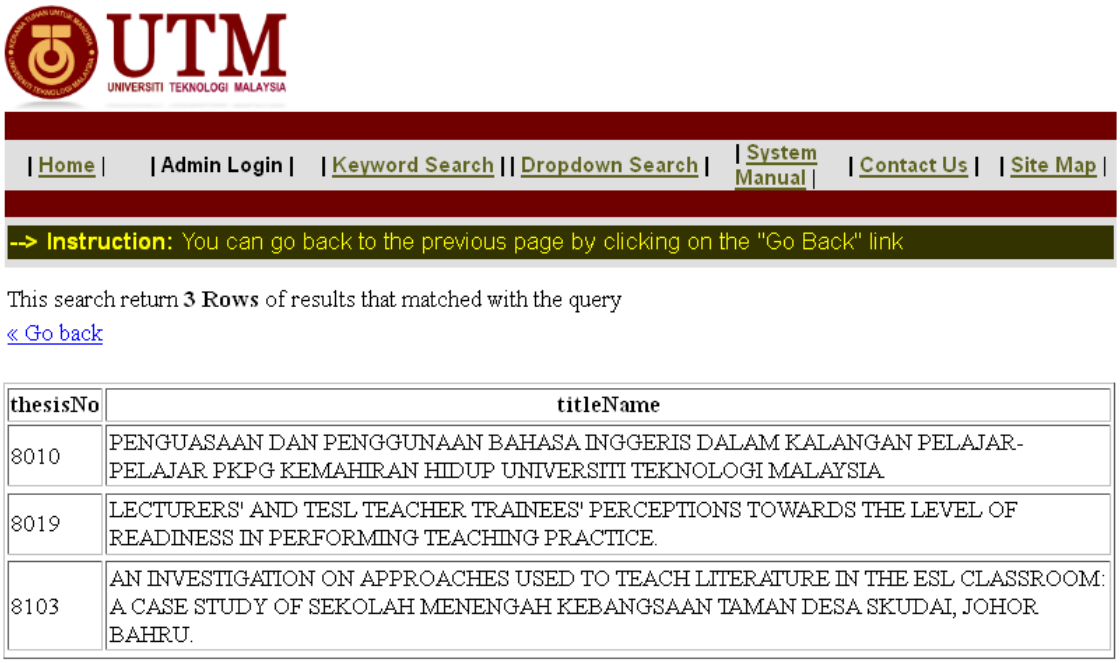


Figure 4.11: Screenshot of expected result for Scenario 3

Table 4.9 shows the choices of combinations of keywords used by the students. The result shows that 30.56% of respondents select the keyword *"field of study"* as their first choice of term followed by *"field of study → subject or course"* (13.86%) and finish with level 3 keyword of *"field of study → subject or course → tesl"* (5.56%). Appendix D shows sample outputs of different keyword combinations for Scenarios 1, 2 and 3.

Table 4.9: List of hierarchy of keywords used in Scenario 3

Row Labels	Data Number of Search Attempts	Percent (%)
education level	**5**	**6.94%**
field of study	**22**	**30.56%**
-	**3**	**4.17%**
-	3	4.17%
other issue	**9**	**12.50%**
assessment	1	1.39%
career	1	1.39%
e-learning	1	1.39%
language	2	2.78%
performance	1	1.39%
personality and social	1	1.39%
-	2	2.78%
subject or course	**10**	**13.89%**
building construction	1	1.39%
electric electrical	1	1.39%
tesl	4	5.56%
-	4	5.56%
learning	**21**	**29.17%**
others	**8**	**11.11%**
teaching	**16**	**22.22%**
Grand Total	**72**	**100.00%**

note: "-" sign means no option is selected

Table 4.10 shows results of Part C. Questions 1 to 9 are questions related to respondents' experience using the prototype system and questions 10 to 17 are questions related to the system interface. Mean rating of each question show that respondents find the system easy to use (3.75) and the system allows them to search for the thesis in English and Bahasa Malaysia title in one search query (3.84).

Table 4.10: Results of Part C

Number	System Evaluation	Mean	Standard deviation
	Part 1: System usage		
1	The system gives me greater control to find what I need	3.47	0.85
2	The system enables me to search the thesis quickly	3.62	0.90
3	The system returns the search results more accurately	3.57	0.89
4	The system returns the search results containing theses written in English and Bahasa Malaysia	3.84	0.78
5	The drop-down menu makes the searching process easier	3.75	0.92
6	The drop-down menu gives me limited searching to the thesis database	2.56	0.99
7	I need to spend a lot of time to learn how to use the system	3.45	1.11
8	I prefer to use the drop-down menu rather than an open keyword search when I am not sure what topic I am looking for	3.48	1.03
9	I feel very confident in using the system	3.48	0.93
	Part 2: System interface		
10	When I start to use the system I know where to start the searching process	3.04	0.99
11	When I use the system I know how to do the searching	3.22	1.01
12	I can easily follow the search steps to get to the search results	3.41	0.80
13	Using the system requires a lot of mental effort	3.60	0.79
14	I like the layout of the system	3.56	0.88
15	The system navigation features are difficult to use (buttons, scrolls, etc)	2.85	0.93
16	The system provides helpful guidance in performing tasks	3.45	0.79
17	Overall, I found that this system is useful in getting the thesis as needed	3.80	0.92

The above results show that the majority respondents are female. The age group is between 18 to 20 years old which matches the education level of "after matriculation level", which is the age most students in Malaysia finish the matriculation level. Most respondents also have skilled knowledge in computer and Internet usage. This trend is due to students being introduced to computer technology in secondary school in Malaysia.

Evaluation results in Part B show that respondents are able to conduct search query to meet the query result of the scenario. Results show that respondents are able to follow logical thinking in making a query and to find the results. Results also show that the majority of the respondents were selecting the same combination of keywords for the first scenario that asked to search for the thesis title of *"lecturers' and TESL teacher trainees' perceptions towards the level of readiness in performing teaching practice".* The first choice of keyword combination is "*teaching → educator → lecturer"* which reflects the title's theme about the teaching environment and involvement of the lecturer and teacher trainee as educator. These clues result in students selecting the "*teaching"* keyword first.

Similarly in the second scenario *"You are required to do research for your course presentation. You have been assigned a topic related to 'teaching and learning in Mathematics'. You have to use the database system provided to search for those related theses and present the outcome about the choice of topics that have been done by the previous students",* the students use the combination of keywords "*field of study → subject or course → mathematics"* which reflects the theme related to Mathematics as a field of study. Note that *"teaching"* or *"learning"* keywords was not selected by the respondents and the majority of respondents choose to use the keywords *"field of study"* which is reflected in the question that is *"mathematics"*. Nevertheless, there are a number of respondents (15.74%) who select *"teaching"* or *"learning"* as their first choice of keyword.

In the third scenario, students are asked to search for a specific thesis based on the given title of *"An investigation on approaches used to teach literature in the ESL classroom: A case study of Sekolah Menengah Kebangsaan Taman Desa Skudai,*

Johor Bahru". The first choice of keyword combination is "*field of study → subject or course → tesl"* which reflects the theme about teaching in ESL classroom. However, there are a large numbers of students who are also selecting *"learning"* keyword as their first choice of keyword after "field of study".

The results also show that it is possible to only use high level keywords to search, which is not necessary to have all keyword combinations to be submitted. This technique is really helpful when the searchers cannot identify specific keywords to use and they can search using only first level keywords to broaden the search results.

The evaluation results in Part C, shows that the majority of respondents expressed a positive viewpoint on the use of drop-down menus to make the search process easier and quicker, and they do not need to consider which keyword to use. Evaluation results also show that the prototype system is able to return search results in both English and Bahasa Malaysia language. Generally, most of the respondents agreed that the system is useful to obtain the thesis title they need and they can follow the search sequence in achieving the results.

4.7 Comparison of the Ontology-based System Prototype and the Current Database System

Experiments have been conducted to evaluate search outcomes using the proposed ontology-based prototype system and the database system, which is non-ontology based. The experiment was conducted using the same scenarios from the survey. Researcher founds that the search result returned from the non-ontology based can lead to unsuccessful search rather than ontology-based prototype system that can conduct a simple search as well as a complicated search. Table 4.11 shows the comparison of search results:

Table 4.11: Comparison of the ontology-based system prototype and the current database system

Scenario	Results based on exact phrase using current database system	Results based on all words using current database system	Results based on any word using current database system	Results obtained from proposed ontology-based prototype system
Scenario 1: *"Lecturers' and TESL teacher trainees' perceptions towards the level of readiness in performing teaching practice".*	1 record found.	1 record found.	8103 records found.	5 records found.
Scenario 2: *"You are required to do research for your course presentation. You have been assigned a topic related to 'teaching and learning in Mathematics'. You have to use the database system provided to search for those related theses and present the outcome about the choice of topics that have been done by the previous students".*	No record found.	No record found.	No record found. However, when searching the topic separately; 939 records found for 'teaching mathematics' and 1005 records found for 'learning mathematics'.	11 records found.
Scenario 3: *"An investigation on approaches used to teach literature in the ESL classroom: A case study of Sekolah Menengah Kebangsaan Taman Desa Skudai, Johor Bahru".*	No record found.	1 record found. (Exact title included)	Microsoft JET Database Engine error '80004005' Query is too complex.	3 records found.

The results have shown that the ontology-based prototype system is able to perform a more effective search compared to the existing database system, which is non-ontology-based.

4.8 Communication

To communicate results on the proposed framework and prototype system, the following list shows articles that have been published:

"*The 11th International DSI and the 16th APDSI Joint Meeting, Taipei, Taiwan, July 12 – 16, 2011"* - Ontology Based Search Mechanism in Bilingual Database Resource.

"International Business Information Management Association (IBIMA). 16th IBIMA Conference 29 – 30 June 2011" - Ontology Information Retrieval for Academic Resources.

"Proceedings of the 4th International Conference on Information Technology and Multimedia at UNITEN (ICIMU 2008), Malaysia. 17th – 19th November 2008" - Improving the Internet Search Capability by Semantic Technology.

4.9 Conclusion

This chapter describes the application of ontology-based search framework to address information retrieval problems faced by novice researchers. The six steps of design science research process introduced by Ken *et. al.* (2007) have been adapted. The six steps are: (1) problem identification and motivation, (2) objectives of solution, (3) design and development, (4) demonstration, (5) evaluation, and (6) communication. The unique query interface is formulated as ontology using a mind-map to enable a dynamic query to be displayed in a hierarchical structure. The evaluation of the case study has been conducted using three case scenarios.

The results show that respondents found that the use of ontology as a query interface has helped in a query formulation. Results also showed that 60.10% of respondents were able to use keywords for their first choice to make retrieval. The majority of respondents agreed that they could use the prototype system confidently to conduct a search query and the prototype system is useful in finding thesis resources.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

This chapter concludes the thesis presentation. The chapter is organised as follows: Section 5.1 presents research findings. Section 5.2 outlines and discusses research contributions. Section 5.3 describes research limitations. Section 5.4 proposes future research direction and concludes this chapter.

5.1 Research Findings

This research has reviewed information retrieval technologies. The research has also reviewed the capability of Semantic Web technologies to improve the information retrieval process. Semantic Web allows data sharing across applications and it is able to link unstructured documents into a web of data to allow data to be linked semantically, which can result in improved searches and retrieval processes.

A case study has been selected in this research to investigate information retrieval problems faced by students; in particular students who are novice researchers. Issues on information retrieval problems have been identified and have been categorised into five groups: (1) query formulation, (2) lack of user perseverance, (3) terminology and search terms, (4) lack of user experience, and (5) synonyms and homonyms. Query formulation refers to searchers not able to formulate the search query effectively. Searchers can face problems in achieving search results that match with their intention, resulting in a lack of perseverance for users if repetitive search efforts still result in fruitless searches. The problem related to terminology and search terms can result in difficulties in query formulation problems. Lack of user experience can result in search results being returned are not those that were expected. Finally, synonyms and homonyms problems result in searchers needing to put in more effort into the search process. This research proposes an ontology-based search framework to overcome the above problems. The use of ontology provides a domain-specific query formulation and to aid in information retrieval. The elements of ontology can be structured in the form of class/subclass to provide a mechanism to structure queries using broad/narrow search terms.

The ontology is designed using a mind-map to represent the way users formulate queries using keywords. In the mind-map, keywords are organised into hierarchical structure of categories and sub-categories. The mind-map aims to help to define meanings of keywords and its relationships. The mind-map works as the hierarchical structure diagram to allow keywords to be categorised in a directional manner from broader to narrow scope. With this method, the profile of keywords is clearer and easily understood to show relationships of different categories.

A system prototype has been developed to demonstrate the ontology-based search framework. The process in the ontology-based search involves a query issued by the user through the query interface. The query interface was designed based on ontology datastore structure to formulate user queries to match the requested query. The ontology datastore consists of classes/subclasses hierarchical structure that represents keywords/sub-keywords structure in the query interface. Once the query is processed, the output will be sent to the user to review. This process will be repeated if the user is not satisfied with the returned documents or output, otherwise the process stops. This way a dynamic search approach is designed to help users to formulate a query.

A case study approach is used to test system functionality and feasibility. The case study selected is a public local university in Malaysia with a multicultural background consisting of local and international students. The case study focuses on novice researchers who used database search systems to search past theses written in either Bahasa Malaysia or English language. These novice researchers are students who are enrolled in different courses in the Faculty of Education with minimal knowledge and skills in database searching. When searching for the past thesis, novice researchers who are facing problems of how to make a query using the open keyword search will find that the hierarchical structure of keywords feature query-guided option easier to use. In addition, this feature can help the search system to return all related titles to ensure results are complete. Novice researchers with a lack of perseverance and lack of search skills will find the structured-terms expressed in the categories-subcategories structure will help them with queries. Furthermore, the

categories-subcategories structure enables novice researchers in determining appropriate keywords to use in the synonym-homonym structure.

This research applies design science research methodology by Ken *et al.* (2007) to assist in prototype development. The methodology practices six steps of the design science research process which consists of: (1) problem identification and motivation, (2) objectives of solution, (3) design and development, (4) demonstration, (5) evaluation, and (6) communication. The steps of design science research methodology has been identified, practiced and clarified in detail when applying the ontology-based search framework to the case study. Each of the steps defined the process in each level of ontology-based search prototype development, to ascertain that the needs meet the research objectives.

The first step of problem identification and motivation begins with identifying the problems faced by novice researchers such as university students who embark on a research programme for the first time. Students often find identifying a research topic to be a complex process due to their inexperience, lack of skill in academic search, or limited literature search experience. Problems such as difficulty in formulating the right keyword in a search query; search results often not matching with expected search outcomes; search results do not match with desired topic area; having difficulty in cross-disciplinary topic searches and searches that can only be performed in mono-language queries are identified.

In the second step of objectives of solution, this research develops an ontology-based search framework to help novice researchers in query formulation to overcome problems identified in step 1. The proposed framework aims to overcome other identified problems, which include synonyms, short-form terms and bilingual queries.

In the third step of design and development, the proposed framework is applied to the thesis database system from the Faculty of Education, UTM. An inductive approach of ontology design is applied. The inductive approach creates ontology by observing, examining and analysing the sample domain. A mind-map expressed in the

hierarchical structure of concepts is used as a basis for the ontology-based structure. In the fourth step of the demonstration, examples of searches that allow shortened forms of terms, synonyms and bilingual searches have been demonstrated through three case scenarios. In the fifth step of evaluation, the prototype system has been evaluated by a group of first year students from the Faculty of Education. The evaluation process aims to investigate system performance and to obtain feedback on whether the proposed system has met the objectives identified. The results have shown that respondents have indicated that the use of ontology as a query interface has helped in the query formulation. The majority of respondents have agreed that they can use the prototype system confidently to conduct search a query and the prototype system is useful in finding thesis resources. Finally, in step 6 (communication), the outcome from this has been published in conferences and in a journal.

5.2 Limitation of Research

This research has its limitations in that the prototype evaluation is restricted to a case study. The ontology datastore mapping process is conducted using an existing thesis database. The bilingual search feature is limited to two languages only that are English and Bahasa Malaysia. The questionnaires used in this research can be extended to be able to identify particular aspects of respondents' views and findings, which can help in upgrading the system.

5.3 Research Contribution

This research has proposed an ontology-based search approach for information retrieval to improve the capability of searches using a dynamic query interface formulated from a mind-map. The mind-map can be designed to mirror the way humans formulate a query. The structure of ontology in a hierarchical manner can be easily managed and properties of ontology can be used to associate properties and relationships such as synonyms. The use of a mind-map to structure ontology is an

innovative contribution to aid in a query formulation to help novice researchers in information retrieval.

In addition, dynamic ontology-based search helps users to identify the most appropriate keyword combinations to formulate the query. The hierarchical structure of ontology offers a way to formulate keyword combinations based on broad and narrow topic areas to search. This innovative approach will be most beneficial to novice researchers when they conduct an initial investigation of possible research topic areas.

5.4 Future Research

The capability of the ontology-based search approach should not focus on dynamical data structure only. Results obtained from this research have shown that the ontology approach is capable of a bilingual search. This aspect can be extended to a multi-lingual search and to investigate properties of ontology that can be utilised for a multi-lingual search.

In addition, future research can be conducted to automatically update the datastore when new data is identified. Currently, the update process is performed manually. Automatic or semi-automatic update process can be investigated in a situation where multiple datastores are to be included in the search database, such as from multiple data sources.

Future research can also be conducted to enable user-defined mind-maps that can be translated to ontology and be formulated into a dynamic query interface that can be personalised or customised. This capability will offer an individualised query user interface to suit the needs of individual users.

This research can be extended by comparing with other ontology-based search framework, for example the Methontology (Lopez *et al.*, 1999) and the approach proposed by Bernaras *et al.* (1996). The Methontology framework enables the construction of ontology at the knowledge level which has been used in the Chemical

OntoAgent, Ontogeneration and $(Onto)^2$ Agent application. The approach proposed by Bernaras *et al.* (1996) needs new ontology to be built that represents the knowledge required by reusing existing ontology or integrated into previous ontology and the approach has been used in the domain of electrical network.

REFERENCES

Abusalah, M., Tait, J. and Oakes, M. (2009) Cross Language Information Retrieval Using Multilingual Ontology as Translation and Query Expansion Base. *Polibits,* 40, 13-16.

Aggarwal, S., Hung, F. and Meng, W. (1998) WIRE-A WWW-Based Information Retrieval and Extraction System. *9th International Conference and Workshop on Database and Expert Systems Applications (DEXA'98),* 24-28 August 1998, Vienna, Austria**,** 887-892.

Al-Maskari, A. and Sanderson, M. (2011) The Effect of User Characteristics on Search Effectiveness in Information Retrieval. *Information Processing & Management,* 47(5), 719-729. DOI:10.1016/j.ipm.2011.03.002.

Antell, K. and Jie, H. (2008) Subject Searching Success Transaction Logs, Patron Perceptions, and Implications for Library Instruction. *Reference & User Services Quarterly,* 48 (1), 68-76.

Arpírez, J. C., Corcho, O., Fernández-López, M. and Gómez-Pérez, A. (2001) Webode: A Scalable Workbench for Ontological Engineering. *Proceeding of the 1st International Conference on Knowledge Capture*, Victoria, British Columbia, Canada**,** 6-13.

Baeza-Yates, R. and Ribeiro-Neto, B. (1999) *Modern Information Retrieval,* New York: ACM Press.

Bangalore, N. S., Henderson, C. L., Weller, A. C., Blecic, D. D., Dorsch, J. L. and Koenig, M. H. (1998) Using Transaction Log Analysis to Improve OPAC Retrieval Results. *College & Research Libraries*, 59 (1), 39-50.

Bates, M. J. (1989) The Design of Browsing and Berrypicking Techniques for the Online Search Interface. *Online Review,* 13 (5), 17.

Beall, J. (2007) Search Fatigue: Finding a Cure for the Database Blues. *American Libraries,* 38 (3), 46-50.

Beall, J. (2008) The Weaknesses of Full-Text Searching. *The Journal of Academic Librarianship,* 34 (5), 438-444.

Beall, J. and Kafadar, K. (2008) Measuring the Extent of the Synonym Problem in Full-Text Searching. *Evidence Based Library and Information Practice,* 3 (4), 18-33.

Bechhofer, S., Horrocks, I., Goble, C. and Stevens, R. (2001) OilEd: A Reason-Able Ontology Editor for the Semantic Web. *KI 2001: Advances in Artificial Intelligence.* Springer Berlin/Heidelberg.

Becker, J. and Hayes, R. M. (1963) *Information Storage and Retrieval: Tools, Elements, Theories,* New York: Wiley.

Becker, N. J. (2003) Google in Perspective: Understanding and Enhancing Student Search Skills. *New Review of Academic Librarianship,* 9 (1), 84-99. DOI:10.1080/13614530410001692059.

Bernaras, A., Laresgoiti, I. and Corera, J. (1996) Building and Reusing Ontologies for Electrical Network Applications'. *ECAI,* 298-302.

Berners-Lee, T., Cailliau, R., Luotonen, A., Nielson, H. F. and Secret, A. (1994) The World-Wide Web. *Communications of the ACM,* 37 (8), 76-82.

Berners-Lee, T., Hendler, J. and Lassila, O. (2001) The Semantic Web. *Scientific American*.

Birkhofer, H. (2011) From Design Practice to Design Science: The Evolution of a Career in Design Methodology Research. *Journal of Engineering Design,* 2 (5), 333-359. DOI:10.1080/09544828.2011.555392.

Borst, W. N. (1997) Construction of Engineeering Ontologies for Knowledge Sharing and Reuse. University of Twente.

Bozsak, E., Ehrig, M., Handschuh, S., Hotho, A., Maedche, A., Motik, B., Oberle, D., Schmitz, C., Staab, S., Stojanovic, L., Stojanovic, N., Studer, R., Stumme, G., Sure, Y., Tane, J., Volz, R. and Zacharias, V. (2002) Kaon - Towards a

Large Scale Semantic Web. *Third International Conference on E-Commerce and Web Technologies*, London, UK**,** 304 - 313.

Broadbent, R. E., Saunders, G. S. and Ekstrom, J. J. (2006) An Infrastructure for the Evaluation and Comparison of Information Retrieval Systems. *SIGITE'06.* Minneapolis, Minnesota, USA.

Brusa, G., Laura Caliusco, M. and Chiotti, O. (2008) Towards Ontological Engineering: A Process for Building a Domain Ontology from Scratch in Public Administration. *Expert Systems,* 25 (5), 484-503. DOI:10.1111/j.1468-0394.2008.00471.x.

Cardoso, J. and Sheth, A. P. (2006) The Semantic Web and Its Applications. IN Cardoso, J. & Sheth, A. P. (Eds.) *Semantic Web Services, Processes and Applications.* New York, USA, Springer.

Cascia, M. L., Sethi, S. and Sclaroff, S. (1998) Combining Textual and Visual Cues for Content-Based Image Retrieval on the World Wide Web. *Proceedings of the IEEE Workshop on Content - Based Access of Image and Video Libraries.* IEEE Computer Society.

Chen, H., Wigand, R. T. and Nilan, M. S. (1999) Optimal Experience of Web Activities. *Computers in Human Behavior,* 15 (5), 585-608.

Chen, S.-M. and Horng, Y.-J. (1999) Fuzzy Query Processing for Document Retrieval Based on Extended Fuzzy Concept Networks. *IEEE Transactions on System, Man, and Cybernetics - Part B,* 29 (1).

Chen, S.-M. and Wang, J.-Y. (1995) Document Retrieval Using Knowledge-Based Fuzzy Information Retrieval Techniques. *IEEE Transactions on Systems, Man and Cybernetics,* 25 (5).

Chen, S.-Y. and Rieh, S. Y. (2009) Take Your Time First, Time Your Search Later: How College Students Perceive Time in Web Searching. *Proceedings of the American Society for Information Science and Technology,* 46 (1), 1-19. DOI:10.1002/meet.2009.1450460253.

Chiang, R. H. L., Chua, C. E. H. and Storey, V. C. (2001) A Smart Web Query Method for Semantic Retrieval of Web Data. *Data & Knowledge Engineering*, 38 (1), 63-84.

Chowdhury, G. G. (2004) *Introduction to Modern Information Retrieval*, London: Facet Publishing.

Chowdhury, S., Gibb, F. and Landoni, M. (2010) Uncertainty in Information Seeking and Retrieval: A Study in an Academic Environment. *Information Processing & Management*, 47 (2), 157-175.

Chu, H. (2003) *Information Representation and Retrieval in the Digital Age*, Medford, N.J.: Published for the American Society for Information Science and Technology by Information Today.

Colomb, R. M. (2002) *Information Spaces: The Architecture of Cyberspace*, New York: Springer.

Corcho, O., Fernandez-Lopez, M. and Gomez-Perez, A. (2003a) Methodologies, Tools and Languages for Building Ontologies: Where Is Their Meeting Point? *Data & Knowledge Engineering*, 46(1), 41-64.

Corcho, O. and Gómez-Pérez, A. (2004) Ontology Translation Approaches for Interoperability: A Case Study with Protégé-2000 and WebODE. *The 14th International Conference on Knowledge Engineering and Knowledge Management - EKAW'04*, Northamptonshire, UK**,** 16.

Corcho, Ó., Gómez-Pérez, A., Guerrero-Rodríguez, D. J., Pérez-Rey, D., Ruiz-Cristina, A., Sastre-Toral, T. and Suárez-Figueroa, M. C. (2003b) Evaluation Experiment of Ontology Tools' Interoperability with the WebODE Ontology Engineering Workbench. *The 2nd International Semantic Web Conference (ISWC'03)*, Sanibel Island, Florida, USA, 87.

d'Amato, C., Staab, S., Fanizzi, N. and Esposito, F. (2010) Efficient Resource Retrieval from Semantic Knowledge Bases. *Proceedings of the 2010 IEEE*

Fourth International Conference on Semantic Computing (ICSC), 22-24 Sept. 2010, Pittsburgh, PA**,** 244-251.

Davies, J., Weeks, R. and Krohn, U. (2004) Quizrdf: Search Technology for the Semantic Web. *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*.

Dinet, J., Favart, M. and Passerault, J.-M. (2004) Searching for Information in an Online Public Access Catalogue (OPAC): The Impacts of Information Search Expertise on the Use of Boolean Operators. *Journal of Computer Assisted Learning*, 20 (5), 338-346. DOI:10.1111/j.1365-2729.2004.00093.x.

Ding, L., Finin, T., Joshi, A., Pan, R., Cost, R. S., Peng, Y., Reddivari, P., Doshi, V. and Sachs, J. (2004) Swoogle: A Search and Metadata Engine for the Semantic Web. *Proceedings of the Thirteenth ACM International Conference on Information and Knowledge Management*.

Doan, A., Madhavan, J., Domingos, P. and Halevy, A. (2002) Learning to Map between Ontologies on the Semantic Web. *WWW2002*, Honolulu, Hawaii, USA**,** May 7-11, 2002.

Domingue, J. (1998) Tadzebao and Webonto: Discussing, Browsing, and Editing Ontologies on the Web. *11th Knowledge Acquisition for Knowledge-Based Systems Workshop*.

Dou, D., McDermott, D. and Qi, P. (2005) Ontology Translation on the Semantic Web. IN Stefano, S., Elisa, B., Sushil, J., Roger, K. & Dennis, M. (Eds.) *Journal on Data Semantics Ii*. Springer-Verlag.

Farquhar, A., Fikes, R. and Rice, J. (1997a) The Ontolingua Server: A Tool for Collaborative Ontology Construction. *International Journal of Human-Computer Studies*, 46 (6), 20.

Farquhar, A., Fikes, R. and Rice, J. (1997b) Tools for Assembling Modular Ontologies in Ontolingua. *The Fourteenth National Conference on Artificial Intelligence (AAAI'97)*, Providence, Rhode Island.

Fensel, D. (2000) The Semantic Web and Its Languages. *IEEE Intelligent Systems,* 15(6), 67-73.

Fernandez-Lopez, M., Gomez-Perez, A. and Juristo, N. (1997) Methontology: From Ontological Art Towards Ontological Engineering. *Proceedings of the AAAI97 Spring Symposium* Stanford, USA.

Finin, T., Mayfield, J., Joshi, A., Cost, R. S. and Fink, C. (2005) Information Retrieval and the Semantic Web. *Proceedings of the 38th International Conference on System Sciences*.

Gerber, A. J. (2006) Towards a Comprehensive Functional Layered Architecture for the Semantic Web. University of South Africa.

Griffiths, J. R. and Brophy, P. (2005) Student Searching Behavior and the Web: Use of Academic Resources and Google. *Library Trends,* 53 (4), 539.

Hancock-Beaulieu, M. (1990) Evaluating the Impact of an Online Library Catalogue on Subject Searching Behaviour at the Catalogue and the Shelves. *Journal of Documentation,* Vol. 46 (Iss: 4), pp: 318-338

Hancock-Beaulieu, M. (1993) A Comparative Transaction Log Analysis of Browsing and Search Formulation in Online Catalogues. *Program: Electronic Library and Information Systems,* 27 (3), 269 - 280.

Heijst, G. V., Falasconi, S., Abu-Hanna, A., Schreiber, G. and Stefanelli, M. (1995) A Case Study in Ontology Library Construction. *Artificial Intelligence in Medicine,* 7 39.

Hendler, J. (2005) Knowledge Is Power: A View from the Semantic Web. *AI Magazine,* 26 (4), 76.

Henzinger, M. (2007) Search Technologies for the Internet. *Science,* 317 (5837), 468-471. DOI:10.1126/science.1126557.

Hevner, A. R., March, S. T., Park, J. and Ram, S. (2004) Design Science in Information Systems Research. *MIS Quarterly,* 28 (1), 75-105.

Holsapple, C. W. and Joshi, K. D. (2002) A Collaborative Approach to Ontology Design. *Commun. ACM,* 45 (2), 42-47. DOI:10.1145/503124.503147.

Hölscher, C. and Strube, G. (2000) Web Search Behavior of Internet Experts and Newbies. *Computer Networks,* 33 (1–6), 337-346. DOI:10.1016/s1389-1286(00)00031-1.

Ilhan, S. and Duru, N. (2006) A Comparative Study for Crisp and Fuzzy Query by a Software Tool. *Soft Computing - A Fusion of Foundations, Methodologies and Applications,* 11 (2), 6.

Jansen, B. J. and Pooch, U. (2001) Web User Studies: A Review and Framework for Future Work. *Journal of the American Society of Information Science and Technology,* 52(3), 235 - 246.

Jansen, B. J. and Spink, A. (2004) How Are We Searching the World Wide Web? A Comparison of Nine Search Engine Transaction Logs. *Information Processing & Management,* 42 (1), 248-263.

Java, A., Nirneburg, S., McShane, M., Finin, T., English, J. and Joshi, A. (2007) Using a Natural Language Understanding System to Generate Semantic Web Content. *International Journal on Semantic Web & Information Systems,* 3 (4), 50-74.

Jones, S., Cunningham, S. J., McNab, R. and Boddie, S. (2000) A Transaction Log Analysis of a Digital Library. *International Journal on Digital Libraries,* 3 (2), 152.

Kagolovsky, Y. and Moehr, J. R. (2003) Terminological Problems in Information Retrieval. *Journal of Medical Systems,* 27 (5), 399-408. DOI:10.1023/a:1025687220609.

Kalfoglou, Y. and Schorlemmer, M. (2003) If-Map: An Ontology-Mapping Method Based on Information-Flow Theory.

Kapetanios, E., Sugumaran, V. and Tanase, D. (2008) A Parametric Linguistics Based Approach for Cross-Lingual Web Querying. *Data & Knowledge Engineering.* 11 September 2007 ed., Elsevier.

Ken, P., Tuure, T., Marcus, R. and Samir, C. (2007) A Design Science Research Methodology for Information Systems Research. *J. Manage. Inf. Syst.*, 24 (3), 45-77. DOI:10.2753/mis0742-1222240302.

Kim, A., Hoffman, L. J. and Martin, C. D. (2002) Building Privacy into the Semantic Web: An Ontology Needed Now. *Semantic Web Workshop.* Hawaii USA.

Kim, D.-W. and Lee, K. H. (2001) A New Fuzzy Information Retrieval System Based on User Preference Model. *The 10th IEEE International Conference on Fuzzy Systems,* 1, 127 - 130

Kim, H.-J., Kim, W. and Kim, J. (2006) An Intelligent Shopping Agent for Optimal Purchasing Decision on the Semantic Web. 192-201.

Kim, T., Sim, C.-M., Yuh, S., Jung, H., Kim, Y.-K., Choi, S.-K., Park, D.-I. and Choi, K. S. (1999) Fromto-Clir[tm]: Web-Based Natural Language Interface for Cross-Language Information Retrieval. *Information Processing and Management.* Elsevier Science Ltd.

Klink, S., Ley, M., Rabbidge, E., Reuther, P., Walter, B. and Weber, A. (2004) Browsing and Visualizing Digital Bibliographic Data. IN O. Deussen, C. H., D.A. Keim, D. Saupe (Ed. *Joint EUROGRAPHICS - IEEE TCVG Symposium on Visualization (2004).*

Kohler, J., Philippi, S., Specht, M. and Ruegg, A. (2006) Ontology Based Text Indexing and Querying for the Semantic Web. *Knowledge-Based Systems,* 19, 10.

Korfhage, R. R. (1997) *Information Storage and Retrieval,* New York: Wiley Computer Publishing.

Kress, N., Del Bosque, D. and Ipri, T. (2011) User Failure to Find Known Library Items. *New Library World,* Vol. 112 (Iss: 3/4,), pp.150 - 170. DOI:10.1108/03074801111117050.

Lange, B., Pezuela, C., Theodorou, I., De Antonio, J. M., Nowak, K., Fluegge, M., Barnickel, N., Graubmann, P. and Osinski, S. (2007) Semantic Interoperability – Report About Technological State-of-the-Art. Creative Common.

Large, A. and Beheshti, J. (1997) OPACs: A Research Review. *Library & Information Science Research,* 19 (2), 111-133. DOI:http://dx.doi.org/10.1016/S0740-8188(97)90039-6.

Large, A., Tedd, L. A. and Hartley, R. J. (1999) *Information Seeking in the Online Age: Principles and Practice,* London: Bowker-Saur.

Lau, E. P. and Goh, D. H.-L. (2006) In Search of Query Patterns: A Case Study of a University OPAC. *Information Processing and Management,* 42, 13.

Levy, A. Y., Srivastava, D. and Kirk, T. (1995) Data Model and Query Evaluation in Global Information Systems. *Journal of Intelligent Information Systems,* 5 (2), 121-143. DOI:10.1007/bf00962627.

Lim, E.-P. and Ng, W. K. (1997) A Relational Interface for Heterogeneous Information Sources. *Proceedings of the IEEE International Forum on Research and Technology Advances in Digital Libraries.* Washington, D.C., United States, IEEE Computer Society.

Lin, F. and Mitamura, T. (2004) Keyword Translation from English to Chinese for Multilingual Qa. IN Frederking, R. & Taylor, K. (Eds.) *Machine Translation: From Real Users to Research.* Springer Berlin / Heidelberg.

Lopez, M. F., Gomez-Perez, A., Sierra, J. P. and Sierra, A. P. (1999) Building a Chemical Ontology Using Methontology and the Ontology Design Environment. *Intelligent Systems and their Applications, IEEE,* 14 (1), 37-46.

Luah, A. K., Ng, W. K., Lim, E. P., Lee, W. P. and Cao, Y. Y. (1999) Locating Web Information Using Web Checkpoints. *Proceedings of the 10th International Workshop on Database and Expert Systems Applications*, 1999**,** 716-720. DOI:10.1109/dexa.1999.795272.

Lynch, C. (1997) The Z39.50 Information Retrieval Standard Part I: A Strategic View of Its Past, Present and Future. *D-Lib Magazine,* 3 (4). DOI:citeulike-article-id:634840.

Maedche, A. and Staab, S. (2001) Comparing Ontologies-Similarity Measures and a Comparison Study. Karlsruhe, Germany, Institute AIFB, University of Karlsruhe.

Malki, M., Benslimane, S., Rahmouni, M. and Rahmoun, A. (2008) Towards Ontology Extraction from Data-Intensive Web Sites: An HTML Forms-Based Reverse Engineering Approach. *International Arab Journal of Information Technology,* 5 (1), 34-44.

March, S. T. and Smith, G. F. (1995) Design and Natural Science Research on Information Technology. *Decision Support Systems,* 15 (4), 251-266.

Marchionini, G., McDonald, S. and Tait, J. (2004) From Information Retrieval to Information Interaction. *Advances in Information Retrieval.* Springer Berlin / Heidelberg.

Marchionini, G. and White, R. (2007) Find What You Need, Understand What You Find. *International Journal of Human-Computer Interaction,* 23 (3), 205-237.

Martin, P., Powley, W. and Weston, A. (1998) Using Metadata to Query Passive Data Sources. *Proceedings of the Thirty-First Hawaii International Conference on System Sciences*, 6-9 Jan 1998, 7**,** 286-294. DOI:10.1109/hicss.1998.649223.

McCallum, A. (1996) Bow: A Toolkit for Statistical Language Modeling, Text Retrieval, Classification and Clustering.

McGuinness, D. L., Fikes, R., Rice, J. and Wilder, S. (2000) The Chimaera Ontology Environment. *American Association for Artificial Intelligence*.

Meadow, C. T., Boyce, B. R., Kraft, D. H. and Barry, C. (2007) *Text Information Retrieval Systems,* Amsterdam, London: Academic Press.

Mena, E., Illarramendi, A., Kashyap, V. and Sheth, A. P. (2000) Observer: An Approach for Query Processing in Global Information Systems Based on Interoperation across Pre-Existing Ontologies. *Distributed and Parallel Databases,* 8 (2), 223-271. DOI:10.1023/a:1008741824956.

Menczer, F. and Belew, R. K. (1998) Adaptive Information Agents in Distributed Textual Environments. *Proceedings of the Second International Conference on Autonomous Agents.* Minnesota, United States, ACM.

Missikoff, M. and Taglino, F. (2003) Symontox: A Web-Ontology Tool for Ebusiness Domain. *Proceedings of the 4th International Conference on Web Information Systems Engineering (WISE2003),* Roma, Italy.

Mohd Zaid, N. and Lau, S. K. (2008) Improving the Internet Search Capability by Semantic Technology. *Proceedings of the 4th International Conference on Information Technology and Multimedia (ICIMU' 2008),* 17th – 19th November 2008, UNITEN Malaysia.

Morupisi, L. and Mooko, N. P. (2006) Using the Online Public Access Catalogue at the University of Botswana. *Information Development,* 22 (3), 197-204.

Novak, J. D. and Cañas, A. J. (2006) The Origins of the Concept Mapping Tool and the Continuing Evolution of the Tool. *Information Visualization,* 5 (3), 175-184. DOI:10.1057/palgrave.ivs.9500126.

Noy, N. F. (2004) Tools for Mapping and Merging Ontologies. IN Staab, S. & Studer, R. (Eds.) *Handbook on Ontologies.* Germany, Springer.

Noy, N. F. and Musen, M. A. (2003) The Prompt Suite: Interactive Tools for Ontology Merging and Mapping. *International Journal of Human-Computer Studies,* 59 (6), 983-1024.

Nyulas, C. and Tu, S. (2007) Datamaster – a Plug-in for Importing Schemas and Data from Relational Databases into Protégé. *In Proceedings of the 10th International Protégé Conference*.

Pass, G., Chowdhury, A. and Torgeson, C. (2006) A Picture of Search. *Proceedings of the 1st International Conference on Scalable Information Systems.* Hong Kong, ACM.

Qasem, A., Dimitrov, D. A. and Heflin, J. (2008) Efficient Selection and Integration of Data Sources for Answering Semantic Web Queries. *Proceedings of the 2nd IEEE International Conference on Semantic Computing (ICSC 2008),* 4-7 August 2008, Santa Clara, California**,** 245-252. DOI:10.1109/ICSC.2008.31.

Quan, D. (2007) Improving Life Sciences Information Retrieval Using Semantic Web Technology. *Briefing in Bioinformatics,* 8 (3), 172-182.

rad, M. P., Hassanpour, H. and Pourshaikh, R. (2010) Concept-Based Information Retrieval with Ontology Approach for Cross-Language Searching. *World Applied Sciences Journal 8 (8),* 965-971.

Robertson, S. E. (1977) The Probability Ranking Principle in IR. *Journal of Documentation,* 33 (4), 294 - 304.

Robertson, S. E., Maron and Cooper, W. S. (1982) Probability of Relevance: A Unification of Two Competing Models for Information Retrieval. *Information Technology: Research and Development,* Vol. 1, pp. 1 -21.

Rubin, D. L., Noy, N. F. and Musen, M. A. (2007) Protégé: A Tool for Managing and Using Terminology in Radiology Applications. *Journal of Digital Imaging,* 20, 12.

Sandra, H. (1995) Semantic Interoperability. *ACM Computing Survey,* 27 (2), 271-273. DOI:10.1145/210376.210392.

Sato, K. (2004) Perspectives of Design Research: In Design: Collective Views for Forming the Foundation of Design Research. *Visible Language,* 38 (2), 218-237.

Savoy, J. (2003) Cross-Language Information Retrieval: Experiments Based on CLEF 2000 Corpora. *Information Processing & Management*, 39, 40.

Song, J.-f., Zhang, W.-m., Xiao, W.-d., Li, G.-h. and Xu, Z.-n. (2005) Ontology-Based Information Retrieval Model for the Semantic Web. *Proceedings of the 2005 IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE'05) on e-Technology, e-Commerce and e-Service.* IEEE Computer Society.

Spink, A., Wolfram, D., Jansen, M. B. J. and Saracevic, T. (2001) Searching the Web: The Public and Their Queries. *Journal of the American Society for Information Science and Technology,* 52 (3), 226-234. DOI:10.1002/1097-4571(2000)9999:9999<::aid-asi1591>3.0.co;2-r.

Steffen, S., Rudi, S., Hans-Peter, S. and York, S. (2001) Knowledge Processes and Ontologies. *IEEE Intelligent Systems*, 16 (1), 26-34. DOI:10.1109/5254.912382.

Straccia, U. and Troncy, R. (2006) Towards Distributed Information Retrieval in the Semantic Web: Query Reformulation Using the OMAP Framework *Proceeding of the 3rd European Semantic Web Conference*, 378-392.

Stumme, G. and Maedche, A. (2001) FCA-MERGE: Bottom-up Merging of Ontologies. Karlsruhe, Germany, Institute AIFB, University of Karlsruhe.

Sure, Y., Erdmann, M., Angele, J., Staab, S., Studer, R. and Wenke, D. (2002) Ontoedit: Collaborative Ontology Development for the Semantic Web. *International Semantic Web Conference - ISWC'02*, Sardinia, Italy**,** 221–235.

Swain, M. J. (1999) Searching for Multimedia on the World Wide Web. *IEEE International Conference on Multimedia Computing and Systems*, Jul 1999, 1**,** 32-37. DOI:10.1109/mmcs.1999.779116.

Swartout, B., Patil, R., Knight, K. and Russ, T. (1997) Toward Distributed Use of Large-Scale Ontologies. *Symposium on Ontological Engineering.* USA, Association for the Advancement of Artificial Intelligence - AAAI.

ter Hofstede, A. H. M., Proper, H. A. and van der Weide, T. P. (1996) Query Formulation as an Information Retrieval Problem. *The Computer Journal,* 39 (4), 255-274. DOI:10.1093/comjnl/39.4.255.

Thomas, R. G. (1993) A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition,* 5 (2), 199-220. DOI:http://dx.doi.org/10.1006/knac.1993.1008.

Thuraisingham, B. M. (2004) Security Standards for the Semantic Web. *Science Direct: Computer Standards & Interfaces,* 27 (2005), 257-268.

Toms, E. G. (2000) Understanding and Facilitating the Browsing of Electronic Text. *International Journal of Human-Computer Studies,* 52 (3), 423-452.

Tudorache, T., Noy, N. F. and Musen, M. A. (2008) Collaborative Protégé: Enabling Community-Based Authoring of Ontologies. *7th International Semantic Web Conference - ISWC'08*, Karlsruhe, Germany.

Uschold, M. and Gruninger, M. (1996) Ontologies: Principles, Methods and Applications. *Knowledge Engineering Review.*

Villén-Rueda, L., Senso, J. A. and de Moya-Anegón, F. (2007) The Use of OPAC in a Large Academic Library: A Transactional Log Analysis Study of Subject Searching. *The Journal of Academic Librarianship,* 33 (3), 327-337.

Wang, X., Fang, F. and Fan, L. (2008a) Ontology-Based Description of Learning Object. *The 7th International Conference on Web-based Learning (ICWL 2008)*, Zhejiang Normal University, Jinhua, China**,** 468 - 476.

Wang, Y.-C., Tsai, T.-H. R. and Hsu, W.-L. (2008b) Web-Based Pattern Learning for Named Entity Translation in Korean-Chinese Cross-Language Information Retrieval. *Expert Systems with Applications*.

Weinberg, B. H. (1995) Why Postcoordination Fails the Searcher. *The Indexer,* 19 (3).

Xie, H. I. (2004) Online Ir System Evaluation: Online Databases Versus Web Search Engines. *Online IR System Evaluation,* 28 (3), 211 - 219. DOI:10.1108/14684520410543652.

Yao, Y. Y. (1995) Measuring Retrieval Effectiveness Based on User Preference of Documents. *Journal of the American Society for Information Science,* 46 (2), 133-145.

Yigal Arens, Craig A. Knoblock and Wei-Min Shen. (1996) Query Reformulation for Dynamic Information Integration. *Journal of Intelligent Information Systems,* Special Issue on Intelligent Information Integration, 6(2/3), 1996, 99-130.

Yilu, Z., Jialun, Q., Hsinchun, C. and Nunamaker, J. F. (2005) Multilingual Web Retrieval: An Experiment on a Multilingual Business Intelligence Portal. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS '05)*, 03-06 Jan. 2005**,** 43a-43a.

Yu, H. and Young, M. (2004) The Impact of Web Search Engines on Subject Searching in OPAC. *Information Technology and Libraries*, 13.

Yu, Y. and Peng, Y. (2007) Learning the Semantic Meaning of a Concept from the Web. *Springer-Verlag*, 98-108.

Yuwono, B. and Lee, D. L. (1996) Wise: A World Wide Web Resource Database System. *IEEE Transactions on Knowledge and Data Engineering,* 8(4), 548-554.

Zhou, B. and Yao, Y. (2009) Evaluating Information Retrieval System Performance Based on User Preference. *Journal of Intelligent Information Systems,* 34 (3), 227-248. DOI:10.1007/s10844-009-0096-5.

APPENDIX A – ONTOLOGY DATASOURCE HIERARCHY

ONTOLOGY DATASOURCE HIERARCHY

1st level	2nd level	3rd level
Teaching	Educator	Lecturer
		Mentor
		PKPG Teacher
		Teacher
		Teacher Trainee
		Technical Teacher
	Course or Subject	Building Construction
		Chemistry
		Computer Science
		Electric Electronic
		Engineering
		Islamic Studies
		Living Skills
		Mathematics
		Physics
		Sports Science
		TESL
	Pedagogy	Self Teaching
	Skill	Communication Skill
		Comprehension Skill
		Data processing Skill
		Defining Skill
		Experimental Design Skill
		Fine Motor Skill
		Forecasting Skill
		Generic Skill
		Gross Motor Skill
		Humanity Skill
		Inference Skill
		Manipulative Skill
		Multiple Intelligence Skill
		Observe Skill
		Pronunciation Skill
		Reading Skill
		Writing Skill
	Style	Grasha Riechmann
	Theory	Behaviourism
		Cognitivism
		Constructivism
	Tool	AutoCAD
		Computer
		E-Learning
		MABC
		Teaching Aid

1st level	2nd level	3rd level
Learning	Learner	ESL Student
		Form 4
		IKBN
		PKPG
		SPACE
		Trainee
		Year 1
		Year 2
		Year 3
		Year 4
	Course or Subject	Building Construction
		Chemistry
		Computer Science
		Electric Electronic
		Engineering
		Islamic Studies
		Living Skills
		Mathematics
		Physics
		Sports Science
		TESL
	Methodology	Computer Assisted Learning
		Computer Based Learning
	Approach	Active
		Collaborative
		Critical Thinking
		Problem Based Learning
		Problem Solving
		Scenario Based Learning
		Self Learning
		Situated Based Learning
	Skill	Communication Skill
		Comprehension Skill
		Data processing Skill
		Defining Skill
		Experimental Design Skill
		Fine Motor Skill
		Forecasting Skill
		Generic Skill
		Gross Motor Skill
		Humanity Skill
		Inference Skill
		Manipulative Skill
		Multiple Intelligence Skill
		Observe Skill
		Pronunciation Skill
		Reading Skill
		Writing Skill

	Strategy	Active Learning
		Case Study
		Role Playing
		Simulation
		Story Telling
	Style	Grasha Riehmann
	Theory	Behaviourism
		Cognitivism
		Constructivism
	Tool	AutoCAD
		Computer
		E-Learning
		MABC
		Teaching Aid
	Room and Area	Canteen
		Classroom
		Convocation
		Laboratory
		Workshop
Field Of Study	Course Or Subject	Building Construction
		Chemistry
		Computer Science
		Electric Electronic
		Engineering
		Islamic Studies
		Living Skills
		Mathematics
		Physics
		Sports Science
		TESL
	Other Issue	Assessment
		Achievement
		Attitude
		Behaviour
		Career
		Language
		Entrepreneurship
		Food Service
		Management
		Motivation
		Rules and Safety
		Social
		Special Education
Education Level	College Level	
	Institute Level	
	Primary School Level	
	Secondary School Level	
	Technical Secondary School Level	
	University Level	

Others	Practical Work	Field Work
		Industrial Training
		Micro Teaching
		Teaching Practice
	Respondent	Aboriginal
		Army
		Athlete
		Children
		Parent
		Race
		Staff
	Room And Area	Canteen
		Classroom
		Convocation
		Laboratory
		Workshop
	Supervision	Academic Supervision
		Teaching Supervision
	Thesis Category	Module Based
		Research Based
		System Based

APPENDIX B – LIST OF BAHASA MALAYSIA TERMS

LIST OF BAHASA MALAYSIA TERMS SHOWN IN MIND-MAP IN FIGURE 4.3 (PAGE 61)

TRANSLATED TO ENGLISH TERMS

Malay	English
Pengajaran	Teaching
Pengajar	Educator
Pedagogi	Pedagogy
Kemahiran	Skill
Gaya	Style
Subjek atau kursus	Subject or Course
Teori	Theory
Peralatan	Tools
AutoCAD	AutoCAD
Komputer	Computer
Bantuan Mengajar	Teaching Aid
MABC	MABC
Pembelajaran dalam talian	E-Learning
Pembelajaran	Learning
Pendekatan	Approach
Subjek atau kursus	Subject or Course
Pembinaan bangunan	Building Construction
Kimia	Chemistry
Sains Komputer	Computer Science
Elektrik Elektronik	Electric Electronic
Kejuruteraan	Engineering
Pendidikan Islam	Islamic Studies
Kemahiran Hidup	Living Skills
Matematik	Mathematics
Fizik	Physics
Sains Sukan	Sports Science
TESL	TESL
Pelajar	Learner
Metodologi	Methodology
Bilik dan kawasan	Room and Area
Kemahiran	Skill
Strategi	Strategy
Gaya	Style
Teori	Theory
Peralatan	Tool
Bidang kajian	Field of Study
Subjek atau kursus	Subject or Course
Isu-isu lain	Other Issues
Penilaian	Assessment
Pencapaian	Achievement
Sikap	Attitude

Tingkah-laku	Behaviour
Kerjaya	Career
Bahasa Inggeris	English Language
Keusahawanan	Entrepreneurship
Servis makanan	Food Service
Pengurusan	Management
Motivasi	Motivation
Peraturan dan keselamatan	Rules and Safety
Sosial	Social
Pendidikan khas	Special Education
Peringkat Pendidikan	Education Level
Peringkat Institut	Institute Level
Peringkat Kolej	College Level
Peringkat Sekolah Rendah	Primary School Level
Peringkat Sekolah Menengah	Secondary School Level
Peringkat Sekolah Menengah Teknikal	Technical Secondary School Level
Peringkat Universiti	University Level
Lain-lain	Others
Responden lain	Other Respondents
Penyeliaan	Supervision
Kerja praktikal	Practical Work
Kerja lapangan	Field Work
Latihan industri	Industrial Training
Pengajran mikro	Micro Teaching
Latihan mengajar	Teaching Practice
Bilik dan kawasan	Room and Area
Kategori tesis	Thesis Category

APPENDIX C1 – SURVEY

University of Wollongong 

PARTICIPATION INFORMATION SHEET

TITLE

Investigation of User Interface and Ease of Use of a Digital Resource Centre System based on Semantic Search Approach

PURPOSE OF THE RESEARCH

This is an invitation to participate in a research study conducted by Norasykin Mohd Zaid and Dr Sim Kim Lau, at the University of Wollongong, Australia. The objective of the study is to examine user interface and ease of use of a Digital Resource Centre System developed by Norasykin Mohd Zaid to help students to search for past thesis in the Education Faculty at Universiti Teknologi Malaysia. The information obtained from this study is extremely useful to help future students to search the thesis database when they begin their research courses in the Honours programs. Specifically the survey will:

- evaluate the system user interface,
- evaluate the effectiveness of the system in terms of accurately returned the search results.

The data collected will be analysed without any identifiable personal information. Only summarised and aggregated data will be presented in the thesis and academic publications.

METHOD AND DEMANDS ON PARTICIPANTS

You will be asked to perform three simple tasks in using the system developed by Norasykin Mohd Zaid. Then, you will be asked to answer a questionnaire anonymously. The entire process of doing the search and answer the questionnaire will take about 30-35 minutes to complete. Typical questions in the questionnaire aim evaluate the user interface and ease of use of the systems. You will be asked to rate in the scale of 1 to 5 where 1 means “strongly disagree” and 5 means “strongly agree” on how user interface features of the systems and the search results you have obtained.

Please note that the questionnaire is returned anonymously with no identifying personal information being recorded. Therefore, you will not be able to withdraw your participation after you have completed the questionnaire because we will not be able to identify your responses.

METHOD OF RETURN THE QUESTIONNAIRE

When you have completed the questionnaire, submit this paper to person incharge.

If you have any questions about this study, please e-mail the investigators Norasykin Mohd Zaid at nmz056@uowmail.edu.au or Sim Kim Lau at simlau@uow.edu.au.

Thank you very much in anticipation of your willingness to participate in this study.

INVESTIGATORS

Norasykin Mohd Zaid
PhD Candidate
School of Information Systems and Technology
University of Wollongong
New South Wales, Australia
Phone: +612-42218159
Email: nmz056@uow.edu.au

Dr Sim Kim Lau
Senior Lecturer and PhD Supervisor
School of Information Systems and Technology
University of Wollongong
New South Wales, Australia
Phone: +612-41214132
Email: simlau@uow.edu.au

How to use this system

Dropdown Search

1. User have to select one of FIVE main categories which are;

   1st level category (Refer to "Keyword Index" menu for complete keywords list)

   Teaching - Anything related to teaching pedagogy, teaching skill, educator, course or subject being taught and etc.

   Learning - Could be any matters, person or stuffs that directly related to learning process (Learning Theory, Learning Approach, Learners and etc).

   Field of Study - Subject or course undertaken by students or taught by teacher/lecturers are being groups under this category.

   Education Level - Some of the researches are done based on certain level of education for example: research done among secondary technical school or college students.

   Others - Most of the keywords which are NOT related to teaching and learning terms have been categorized under others. For example: teaching practice, athletes, academic supervision and etc.

2. User have a choice to select the subcategory of the previous level in the 2nd level of dropdown list OR user can proceed with the submit button.

3. User have a choice to select the subsubcategory of the previous level in the 3rd level of dropdown list OR user can proceed with the submit button.

Example

You would like to find thesis's title about "learning using computer". You are going to select:

1st level of dropdown: *Learning*

2nd level of dropdown: *Tool*

3rd Level of dropdown: *Computer*

Once you submit the form, results should be come out with some of the titles like below:

thesisNo	titleName
7900	KEBERKESANAN PBK BERTAJUK TRIGONOMETRI YANG DISEDIAKAN OLEH KPM BAGI PELAJAR-PELAJAR TINGKATAN 4 DI EMPAT BUAH SEKOLAH DI JOHOR BAHRU.
8031	FAKTOR-FAKTOR YANG MEMPENGARUHI PENGGUNAAN KOMPUTER DI KALANGAN GURU TEKNIKAL DI SEKOLAH MENENGAH TEKNIK BINTULU SARAWAK.

Keyword lists:

1. Teaching

- Educator
- Course or Subject
- Pedagogy
- Skill
- Style
- Theory
- Tool
- Room and Area

2. Learning

- Learner
- Course or Subject
- Methodology
- Approach
- Skill
- Strategy
- Style
- Theory
- Tool
- Room and Area

3. Field of Study

3.1 Course or Subject

- Building Construction
- Chemistry
- Computer Science
- Electric Electronic
- Engineering
- Islamic Studies
- Living Skills
- Mathematics
- Physics
- Sports Science
- TESL

3.2 Other Issue

- Assessment
- Achievement
- Attitude
- Behaviour
- Career
- Language
- Entrepreneurship
- Food Service
- Management
- Motivation
- Rules and Safety
- Social
- Special Education

4. Education Level

-College Level
- Institute Level
- Primary School Level
- Secondary School Level
- Technical Secondary School Level
- University Level

5. Others

5.1 Practical Work

- Field Work
- Industrial Training
- Micro Teaching
- Teaching Practice

5.2 Respondent

- Aboriginal
- Army
- Athlete
- Children
- Parent
- Race
- Staff

5.3 Room and Area

- Canteen
- Classroom
- Convocation
- Laboratory
- Workshop

5.4 Supervision

- Academic Supervision
- Teaching Supervision

5.5 Thesis Category

- Module Based
- Research Based
- System Based

PART A:

Please choose the option that is most relevant to you.

1. I am
    a. Male
    b. Female

2. My age
    a. Under 18
    b. 18-20
    c. 21-23
    d. 24 above

3. My most recent education qualification is
    a. After SPM Intake
    b. After STPM intake
    c. After matriculation intake
    d. Holding a diploma degree

4. My computer skill is
    a. Poor
    b. Below average
    c. Above average
    d. Good
    e. Excellent

5. I have been using Internet for
    a. 0 to 3 months
    b. 3 months to 6 months
    c. 6 months to 1 year
    d. 1 year to 3 years
    e. 3+ years

6. I have used Internet on a
    a. Very rarely used (only sometimes in a year)
    b. Rarely used (only sometimes in a month)
    c. A little often (only sometimes in a week)
    d. Often (everyday)
    e. very often (many times in a day)

PART B:

You have been given an access to the server link as below:

http://web2.fp.utm.my/syikin/index.php

From a given scenario, please answer all questions by referring to the drop-down list search on the system. Depending on your choice of keyword on the system, you may or may not find the query you want.

1. You are required to search for a thesis containing the following thesis's title:

   "LECTURERS' AND TESL TEACHER TRAINEES' PERCEPTIONS TOWARDS THE LEVEL OF READINESS IN PERFORMING TEACHING PRACTICE".

   a) Please write the items that you have used in the drop-down list. Please repeat up to three attempts using different selection from the drop down list items.

No.	1st level	2nd level (if required)	3rd level (if required)	How many records return?	Can you find the given title? **
					Yes/No
1					
2					
3					

   In your search results, can you find the given title? (If the results are too many, you may click ***Edit* *Menu* on the Browser Toolbar then click ***Find*** to search for the title)

   b) Please suggest items for the drop-down list which you think is the best sequence to point to the above title given.

   1st level : _____

   2nd level (if required) : _____

   3rd level (if required) : _____

2. You are required to search for the specific thesis based on the given scenario:

   "You are required to do a research for your course presentation. You have been assigned a topic related to 'teaching and learning in Mathematics'. You have to use the database system provided to search for those related thesis and presents the outcome about the choice of topics that have been done by the previous students".

   a) Please write the items for the drop-down list you have selected up to three attempts.

No.	1st level	2nd level (if required)	3rd level (if required)	How many records return?	Do you think these theses relevant to your search?
					Yes/No
1					
2					
3					

   b) Please suggest the items for the drop-down list which you think are the best sequence that students can easily find the above scenario.

   1st level : _____

   2nd level (if required) : _____

   3rd level (if required) : _____

3. You are required to search for the specific thesis containing the given thesis's title:

   "AN INVESTIGATION ON APPROACHES USED TO TEACH LITERATURE IN THE ESL CLASSROOM: A CASE STUDY OF SEKOLAH MENENGAH KEBANGSAAN TAMAN DESA SKUDAI, JOHOR BAHRU".

   a) Please write the items for the drop-down list you have selected up to three attempts.

No.	1st level	2nd level (if required)	3rd level (if required)	How many records return?	Can you find the given title? ** Yes/No
1					
2					
3					

   In your search results, can you find the given title? (Click ***Edit* *Menu* on the Browser Toolbar then click ***Find*** to search for the title)

   b) Please suggest items for the drop-down list which you think is the best sequence to point to the above title given.

   1st level : _____

   2nd level (if required) : _____

   3rd level (if required) : _____

PART C:

Please tick the most appropriate box where

1	2	3	4	5
Strongly disagree	Disagree	Slightly agree	Agree	Strongly agree

Thinking of the searching process and the search results that you have obtained, please answer the following questions.

No	Items	1	2	3	4	5
1.	The system gives me greater control to find what I need					
2.	The system enables me to search the thesis quickly					
3.	The system returns the search results more accurately					
4.	The system returns the search results containing thesis written in English and Bahasa Malaysia					
5.	The drop-down menu makes the searching process easier					
6.	The drop-down menu gives me limited searching to the thesis database					
7.	I need to spend a lot of time to learn how to use the system					
8.	I prefer to use drop down menu rather than open keyword search when I am not sure what topic I am looking after					
9.	I feel very confident in using the system					

Thinking of the user interface, please answer the following questions.

10.	When I start to use the system I know where to start the searching process					
11.	When I use the system I know how to do the searching					
12.	I can easily follow the search steps to get to the search results					
13.	Using the system requires a lot of mental efforts					
14.	I like the layout of the system					
15.	The system navigation features are difficult to use (buttons, scrolls, etc)					
16.	The system provides helpful guidance in performing tasks					
17.	Overall, I found that this system is useful in getting thesis as needed					

- End of Questionnaires -

APPENDIX C2 – QUESTIONNAIRE PART C (AMENDED)

PART C:

Please tick the most appropriate box where you:

1	2	3	4	5
Strongly disagree	Disagree	Slightly agree	Agree	Strongly agree

Thinking of the thesis searching process and the search results that you have obtained, please answer the following questions:

No	Items	1	2	3	4	5
18.	The system gives me a greater control to find what I need					
19.	The system enables me to search for the thesis quickly					
20.	The system returns the search results more accurately					
21.	The system returns the search results containing the thesis written in English and Bahasa Malaysia					
22.	The drop-down menu makes the searching process easier					
23.	The drop-down menu limits my thesis searching process in the thesis database					
24.	I spend a lot of time to learn on how to use the system					
25.	I prefer to use a drop down menu rather than an open keyword search when I am not sure what is the topic that I am looking for					
26.	I feel very confident in using the system					

Thinking of the user interface, please answer the following questions:

27.	When I start using the system I know where to begin the searching process					
28.	When I use the system I know how to do the searching					
29.	I can easily follow the searching steps to get the search results					
30.	To use the system, it requires a lot of thinking process					
31.	I like the layout of the system					
32.	The system navigational features are difficult to use (buttons, scrolls, etc)					
33.	The system provides helpful guidance in performing tasks					
34.	In overall, I found that this system is useful in searching the thesis					

- End of Questionnaires -

APPENDIX D – SAMPLE SCREEN

SAMPLE SCREENS OF THE TEST SCENARIOS CONDUCTED USING SYSTEM PROTOTYPE

Figure 1: Home page screen

How to use this system

Dropdown Search Page

4. User have to select one of FIVE main categories which are;
   1st level category (Refer to "Keyword Index" menu for complete keywords list)

   Teaching - Anything related to teaching pedagogy, teaching skill, educator, course or subject being taught and etc.

   Learning - Could be any matters, person or stuffs that directly related to learning process (Learning Theory, Learning Approach, Learners and etc).

   Field of Study - Subject or course undertaken by students or taught by teacher/lecturers are being groups under this category.

   Education Level - Some of the researches are done based on certain level of education for example: research done among secondary technical school or college students.

   Others - Most of the keywords which are NOT related to teaching and learning terms have been categorized under others. For example: teaching practice, athletes, academic supervision and etc.

5. User have a choice to select the subcategory of the previous level in the 2nd level of dropdown list OR user can proceed with the submit button.

6. User have a choice to select the sub-subcategory of the previous level in the 3rd level of dropdown list OR user can proceed with the submit button.

Example

You would like to find thesis's title about "learning using computer". You are going to select:

1st level of dropdown: *Learning*

2nd level of dropdown: *Tool*

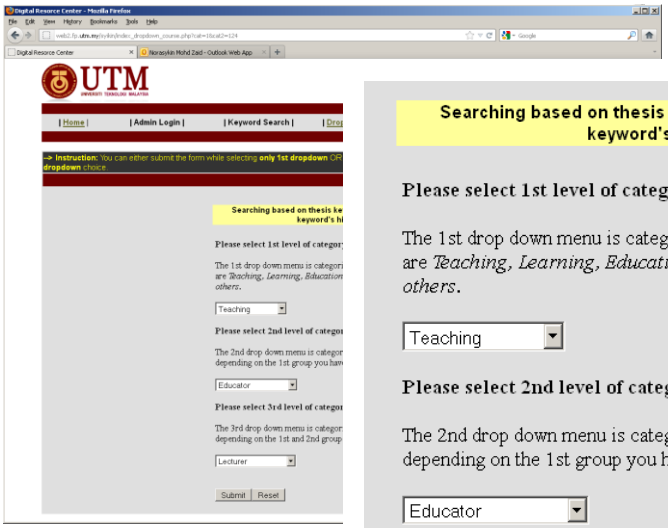
3rd Level of dropdown: *Computer*

Once you submit the form, results should be come out with some of the titles like below:

thesisNo	titleName
7900	KEBERKESANAN PBK BERTAJUK TRIGONOMETRI YANG DISEDIAKAN OLEH KPM BAGI PELAJAR-PELAJAR TINGKATAN 4 DI EMPAT BUAH SEKOLAH DI JOHOR BAHRU.
8031	FAKTOR-FAKTOR YANG MEMPENGARUHI PENGGUNAAN KOMPUTER DI KALANGAN GURU TEKNIKAL DI SEKOLAH MENENGAH TEKNIK BINTULU SARAWAK.

Figure 2: Sample screen showing instruction on how to conduct search using the prototype system

Figure 3: Sample input screen for scenario 1 with the following choice of keywords combination: *Teaching → Educator → Lecturer*



Searching based on thesis keyword listed in dropdown keyword's hierarchy

Please select 1st level of category:

The 1st drop down menu is categorized into 5 major groups which are *Teaching, Learning, Education Level, Field of Study and others.*

Teaching

Please select 2nd level of category (if available):

The 2nd drop down menu is categorized into each subgroups depending on the 1st group you have selected.

Educator

Please select 3rd level of category (if available):

The 3rd drop down menu is categorized into each subsubgroups depending on the 1st and 2nd group you have selected.

Lecturer

Submit Reset

Figure 4: Sample output screen of scenario 1 with the following keywords combination: *Teaching → Educator → Lecturer*

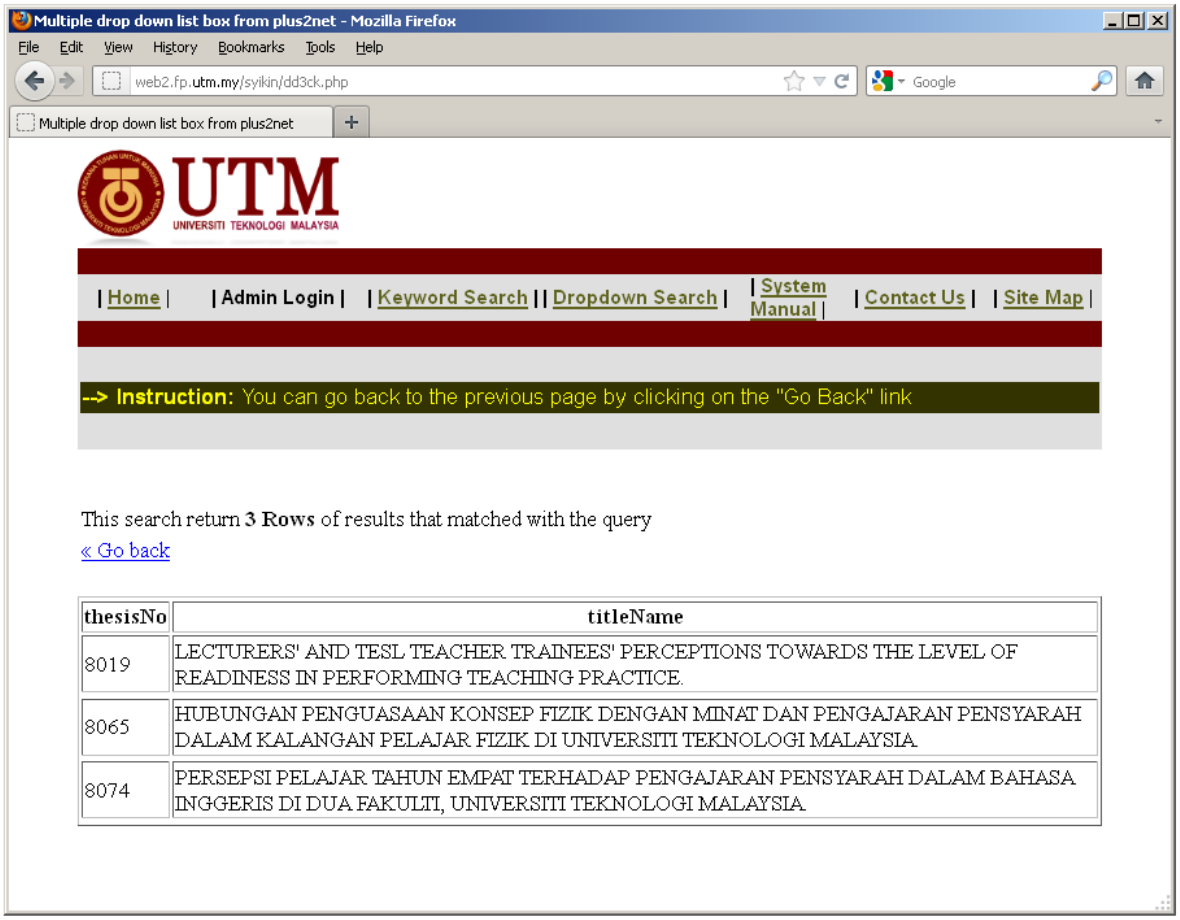


Figure 5: Sample input screen for scenario 1 with the following choice of keywords combination: *Field of study → Subject or Course → TESL*

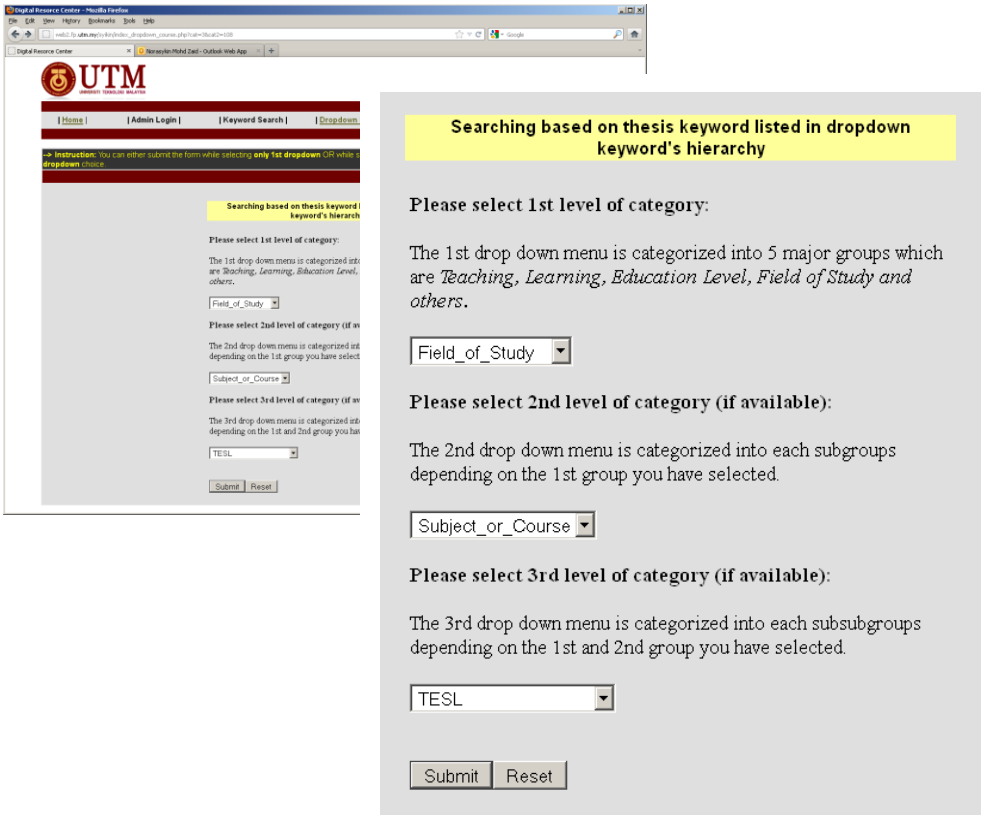


Figure 6: Sample output screen of scenario 1 with the following keywords combination: *Field of study → Subject or Course → TESL*

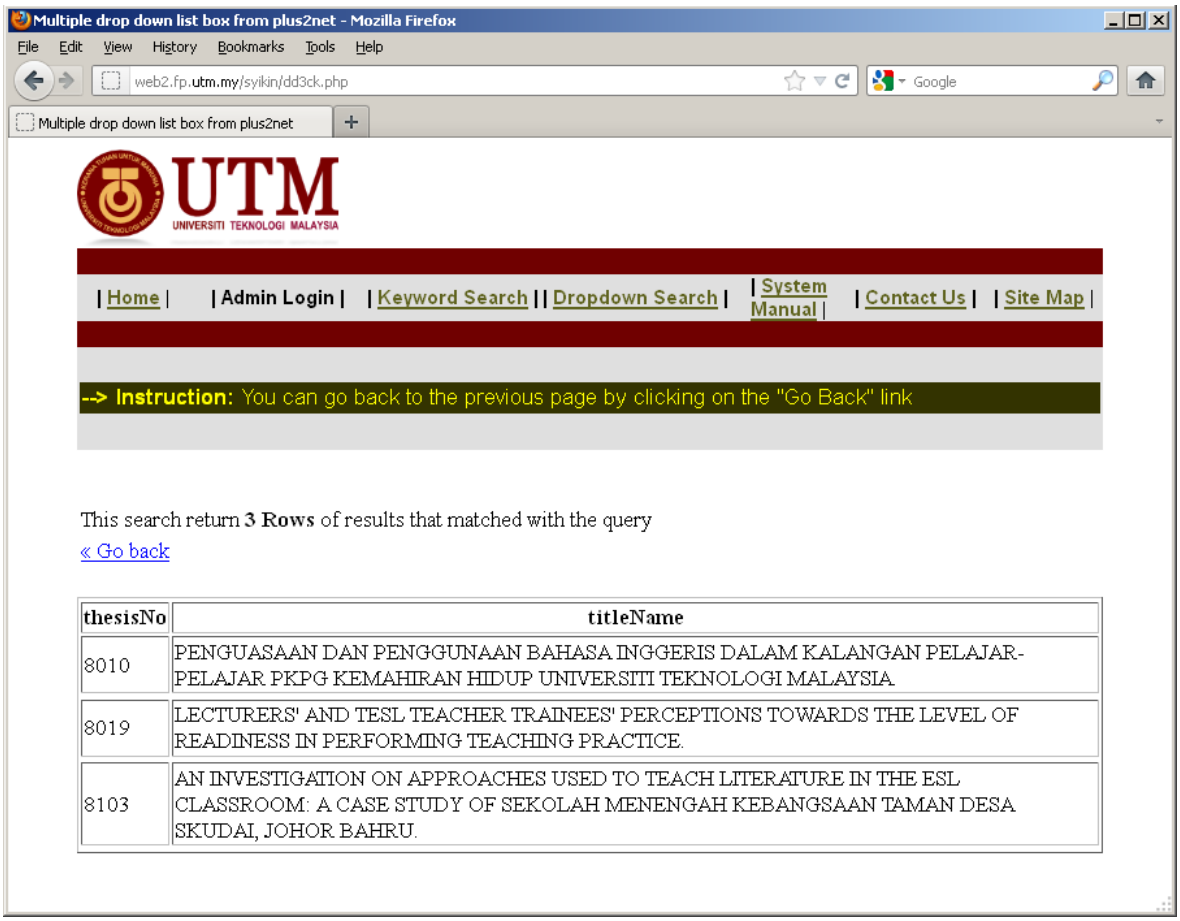


Figure 7: Sample input screen for scenario 1 with the following choice of keywords combination: *Learning → Methodology → NULL*

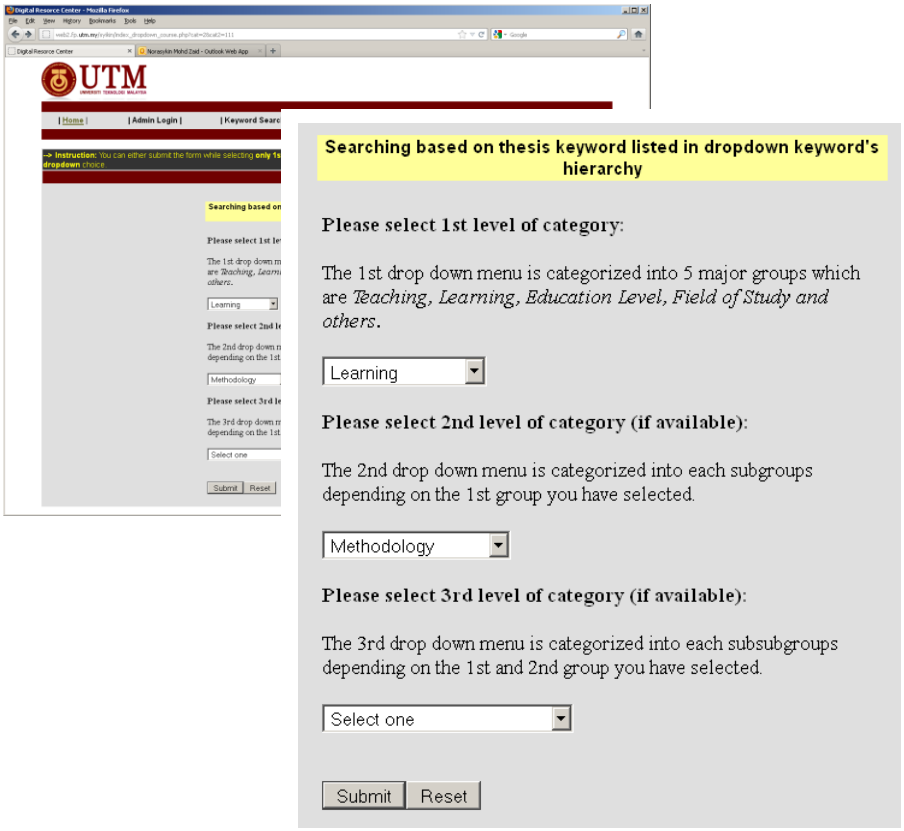


Figure 8: Sample output screen of scenario 1 with the following keywords combination: *Learning → Methodology → NULL*

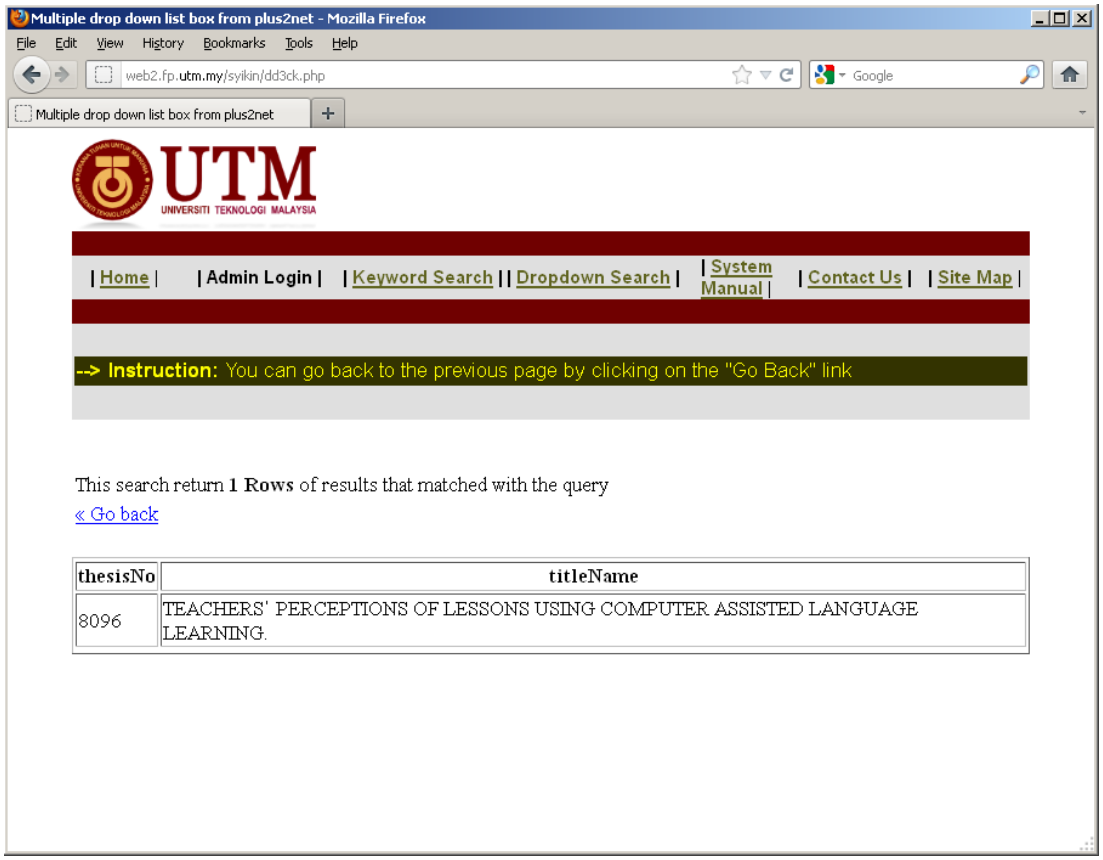


Figure 9: Sample input screen for scenario 2 with the following choice of keywords combination: *Field of Study → Subject or Course → Mathematics*

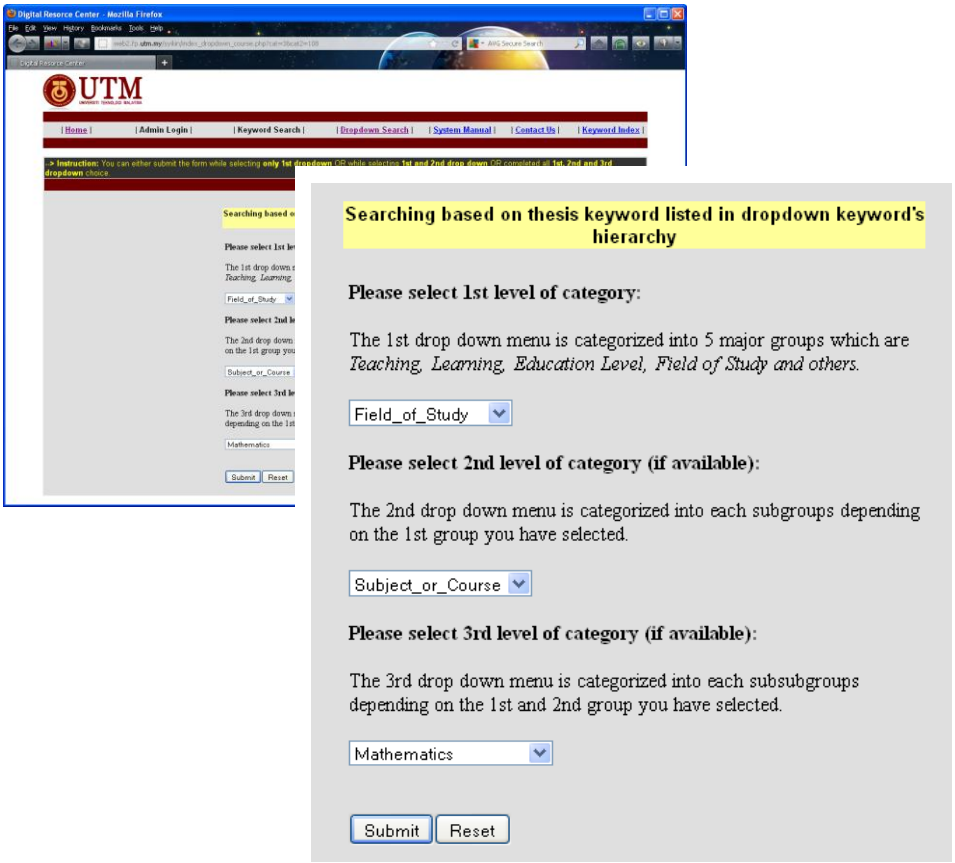


Figure 10: Sample output screen of scenario 2 with the following keywords combination: *Field of Study → Subject or Course → Mathematics*

This search return **11 Rows** of results that matched with the query

« Go back

thesisNo	titleName
8012	PEMBANGUNAN BAHAN E-PEMBELAJARAN BERASASKAN MOODLE BAGI TAJUK STRAIGHT LINE DAN CIRCLES III MATEMATIK TINGKATAN EMPAT.
8023	STRATEGI-STRATEGI PEMAHAMAN DALAM PENYELESAIAN MASALAH MATEMATIK BERPERKATAAN DALAM KURSUS STATISTIK 1.
7900	KEBERKESANAN PBK BERTAJUK TRIGONOMETRI YANG DISEDIAKAN OLEH KPM BAGI PELAJAR-PELAJAR TINGKATAN 4 DI EMPAT BUAH SEKOLAH DI JOHOR BAHRU.
8033	HUBUNGAN ANTARA MINAT PELAJAR DAN SIKAP IBU BAPA DENGAN PRESTASI MATEMATIK TERBAIK PELAJAR PROGRAM SARJANA MUDA SAINS SERTA PENDIDIKAN (MATEMATIK) DAN SARJANA MUDA SAINS DAN KOMPUTER SERTA PENDIDIKAN (MATEMATIK) DI FAKULTI PENDIDIKAN, UNIVERSITI TEKNOLOGI MALAYSIA
8034	PEMBANGUNAN LAMAN WEB BAGI MATAPELAJARAN MATEMATIK KBSM TINGKATAN EMPAT BERTAJUK BULATAN III.
8049	PEMBANGUNAN LAMAN WEB REKREASI MATEMATIK BERASASKAN SUMBER TERBUKA JOOMLA.
8059	MEMBANGUNKAN PERISIAN MODUL BAHAN BANTU MENGAJAR (BBM) BERTAJUK "SOLID GEOMETRY II" BAGI MATA PELAJARAN MATEMATIK TINGKATAN DUA.
8071	PEMBANGUNAN PERISIAN PEMBELAJARAN BERBANTUKAN KOMPUTER (PBK) MATEMATIK TINGKATAN EMPAT BAGI TAJUK LINES AND PLANES IN 3-DIMENSIONS.
7905	PEMBANGUNAN PERISIAN PEMBELAJARAN BERBANTUKAN KOMPUTER (PBK) BAGI TAJUK GARIS LURUS MATEMATIK TINGKATAN EMPAT.
8092	PEMBANGUNAN BAHAN E-PEMBELAJARAN BERASASKAN MOODLE BERTAJUK "QUADRATIC EXPRESSIONS AND EQUATIONS" DAN "STATISTICS" TINGKATAN 4.
8095	PENERAPAN UNSUR SEJARAH DALAM MATEMATIK BAGI TOPIK GEOMETRI KOORDINAT.

Figure 11: Sample input screen for scenario 2 with the following choice of keywords combination: *Learning → NULL → NULL*

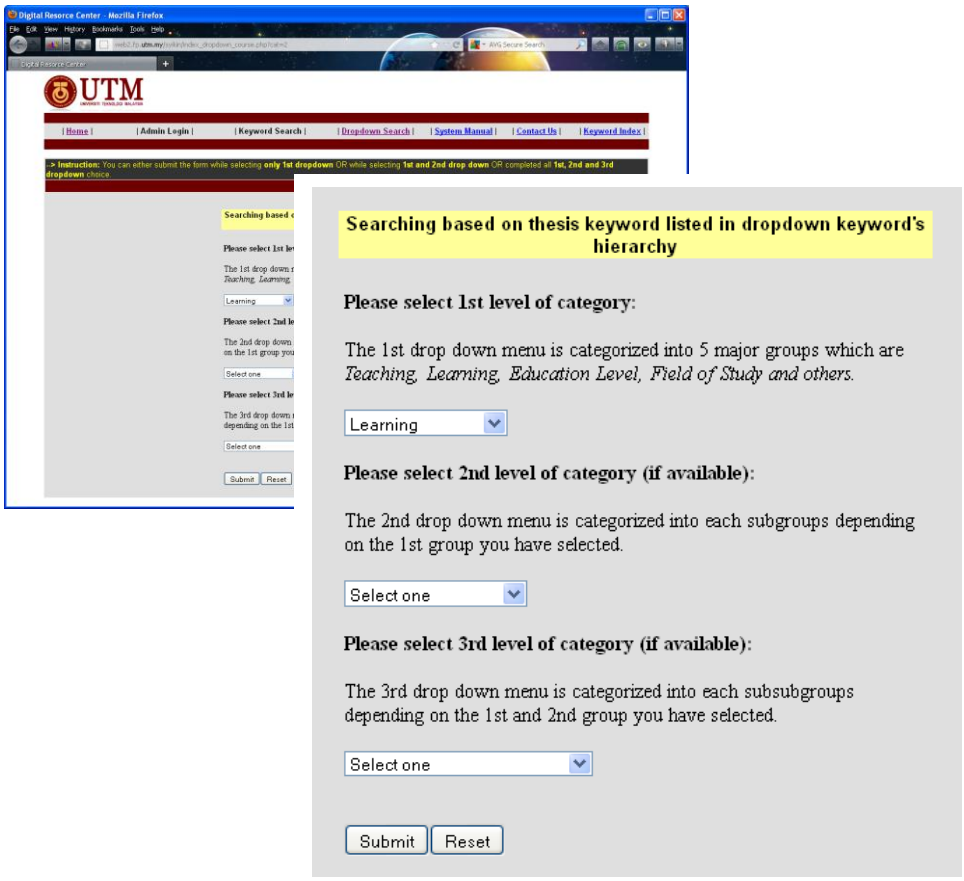
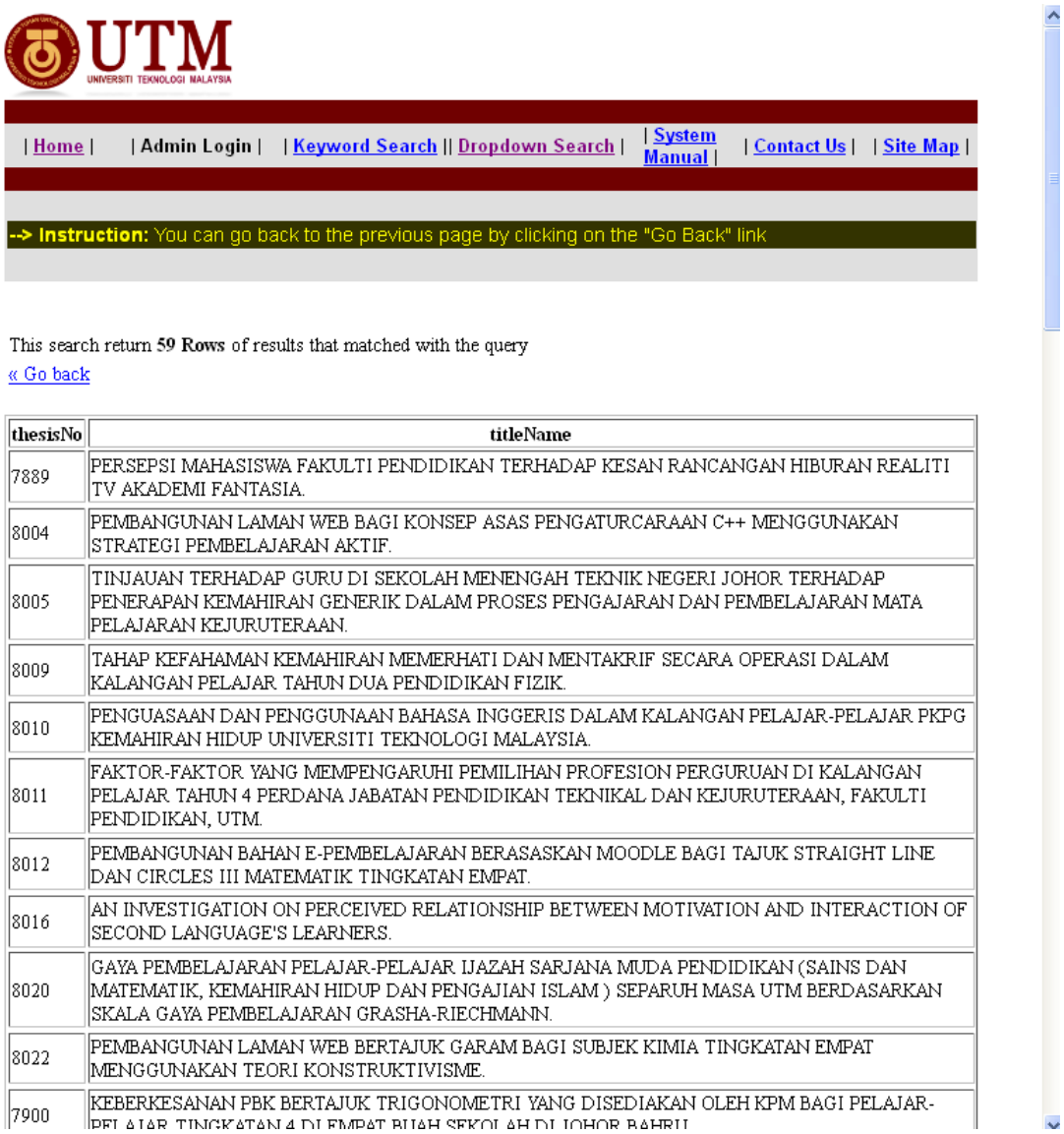


Figure 12: Sample output screen of scenario 2 with the following keywords combination: *Learning → NULL → NULL*



UTM
UNIVERSITI TEKNOLOGI MALAYSIA

| Home | | Admin Login | | Keyword Search || Dropdown Search | | System Manual | | Contact Us | | Site Map |

--> **Instruction:** You can go back to the previous page by clicking on the "Go Back" link

This search return **59 Rows** of results that matched with the query

« Go back

thesisNo	titleName
7889	PERSEPSI MAHASISWA FAKULTI PENDIDIKAN TERHADAP KESAN RANCANGAN HIBURAN REALITI TV AKADEMI FANTASIA.
8004	PEMBANGUNAN LAMAN WEB BAGI KONSEP ASAS PENGATURCARAAN C++ MENGGUNAKAN STRATEGI PEMBELAJARAN AKTIF.
8005	TINJAUAN TERHADAP GURU DI SEKOLAH MENENGAH TEKNIK NEGERI JOHOR TERHADAP PENERAPAN KEMAHIRAN GENERIK DALAM PROSES PENGAJARAN DAN PEMBELAJARAN MATA PELAJARAN KEJURUTERAAN.
8009	TAHAP KEFAHAMAN KEMAHIRAN MEMERHATI DAN MENTAKRIF SECARA OPERASI DALAM KALANGAN PELAJAR TAHUN DUA PENDIDIKAN FIZIK.
8010	PENGUASAAN DAN PENGGUNAAN BAHASA INGGERIS DALAM KALANGAN PELAJAR-PELAJAR PKPG KEMAHIRAN HIDUP UNIVERSITI TEKNOLOGI MALAYSIA.
8011	FAKTOR-FAKTOR YANG MEMPENGARUHI PEMILIHAN PROFESION PERGURUAN DI KALANGAN PELAJAR TAHUN 4 PERDANA JABATAN PENDIDIKAN TEKNIKAL DAN KEJURUTERAAN, FAKULTI PENDIDIKAN, UTM.
8012	PEMBANGUNAN BAHAN E-PEMBELAJARAN BERASASKAN MOODLE BAGI TAJUK STRAIGHT LINE DAN CIRCLES III MATEMATIK TINGKATAN EMPAT.
8016	AN INVESTIGATION ON PERCEIVED RELATIONSHIP BETWEEN MOTIVATION AND INTERACTION OF SECOND LANGUAGE'S LEARNERS.
8020	GAYA PEMBELAJARAN PELAJAR-PELAJAR IJAZAH SARJANA MUDA PENDIDIKAN (SAINS DAN MATEMATIK, KEMAHIRAN HIDUP DAN PENGAJIAN ISLAM) SEPARUH MASA UTM BERDASARKAN SKALA GAYA PEMBELAJARAN GRASHA-RIECHMANN.
8022	PEMBANGUNAN LAMAN WEB BERTAJUK GARAM BAGI SUBJEK KIMIA TINGKATAN EMPAT MENGGUNAKAN TEORI KONSTRUKTIVISME.
7900	KEBERKESANAN PBK BERTAJUK TRIGONOMETRI YANG DISEDIAKAN OLEH KPM BAGI PELAJAR-PELAJAR TINGKATAN 4 DI EMPAT BUAH SEKOLAH DI JOHOR BAHRU

Figure 13: Sample input screen for scenario 2 with the following choice of keywords combination:

Teaching → Educator → Lecturer

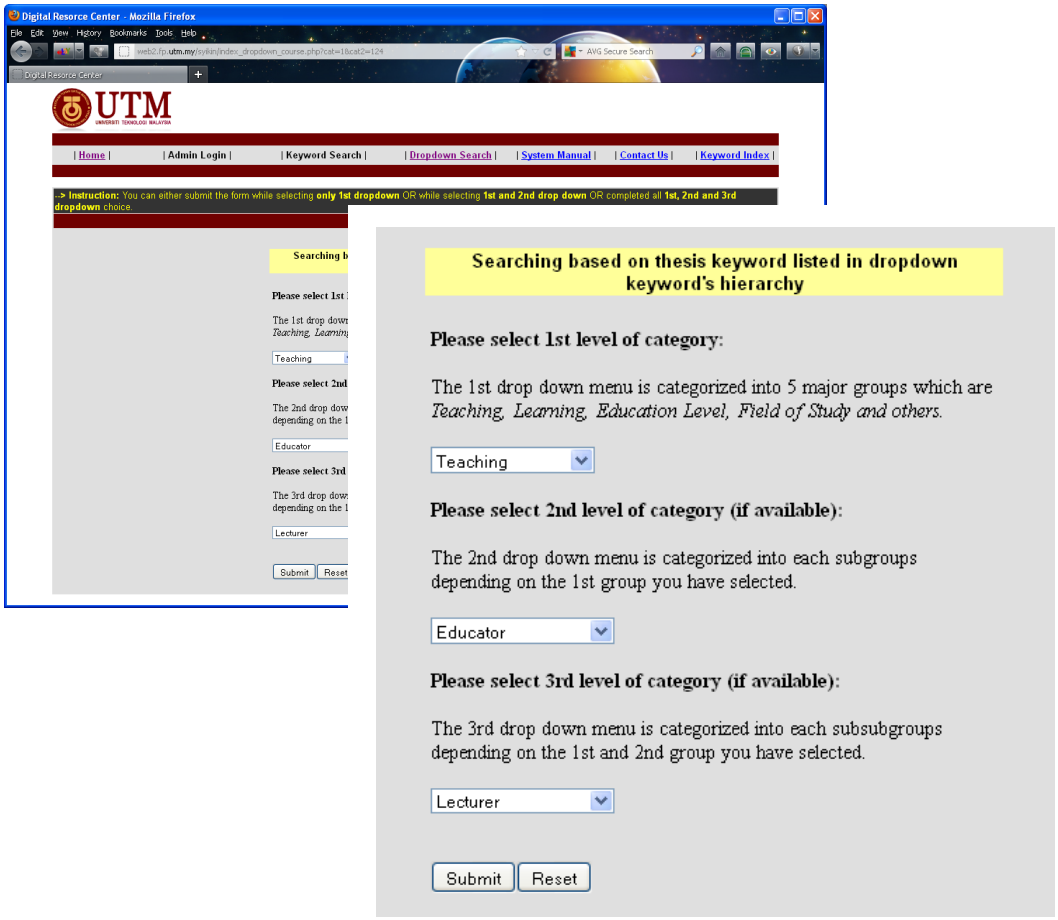


Figure 14: Sample output screen of scenario 2 with the following keywords combination: *Teaching → Educator → Lecturer*

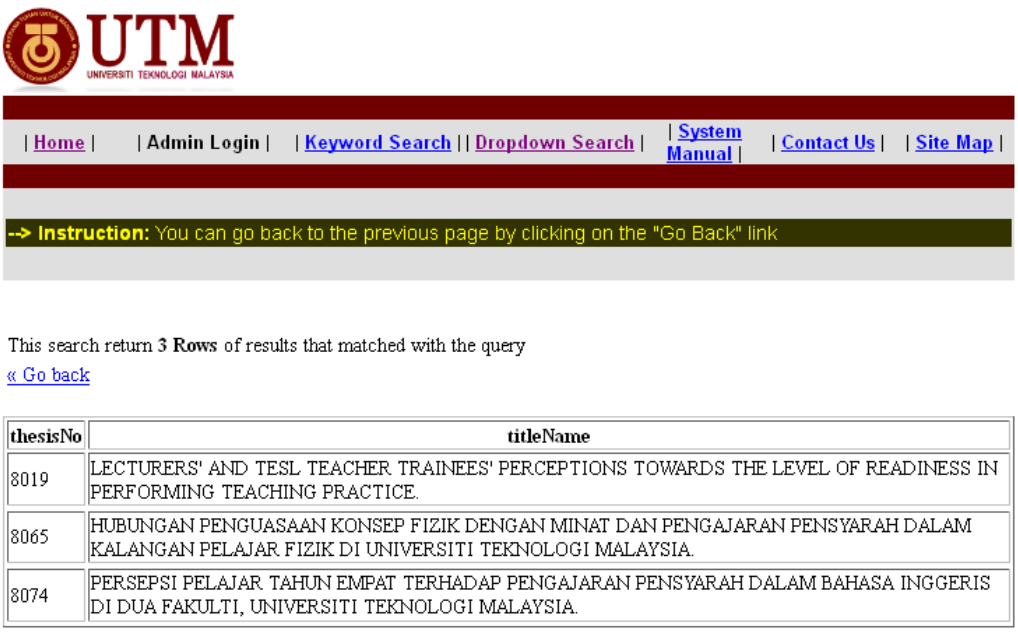


Figure 15: Sample input screen for scenario 3 with the following choice of keywords combination: *Field of Study → Subject or Course → TESL*

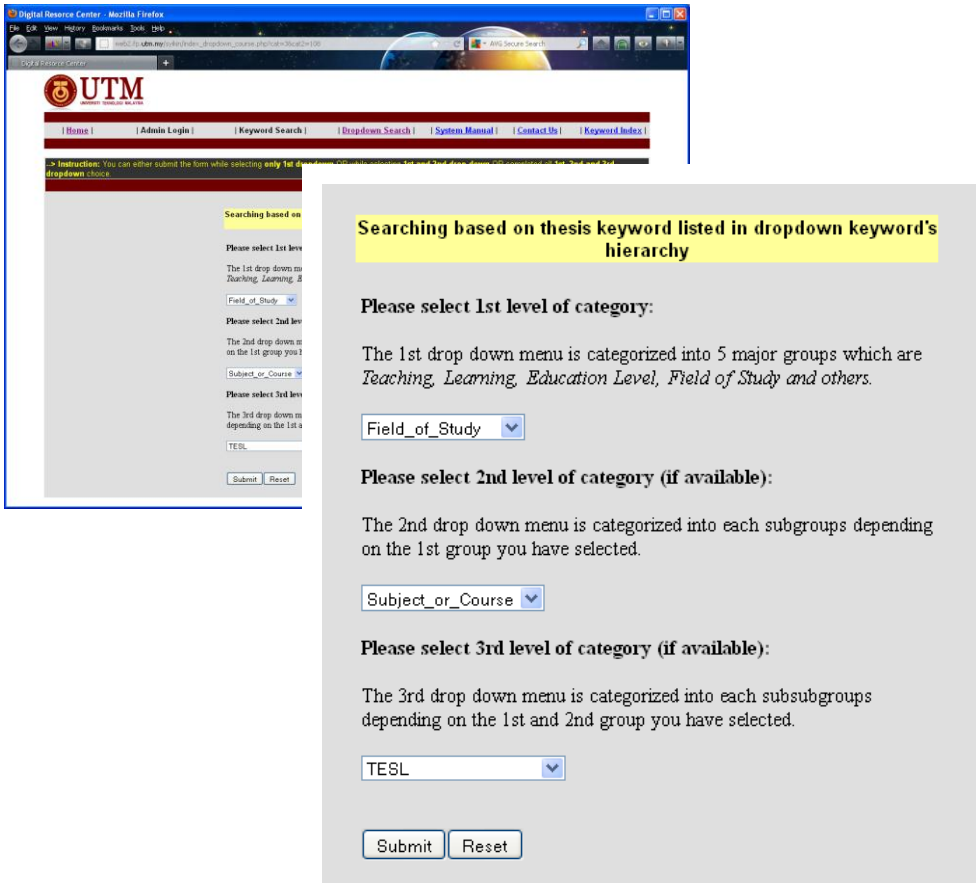


Figure 16: Sample output screen of scenario 3 with the following keywords combination: *Field of Study → Subject or Course → TESL*

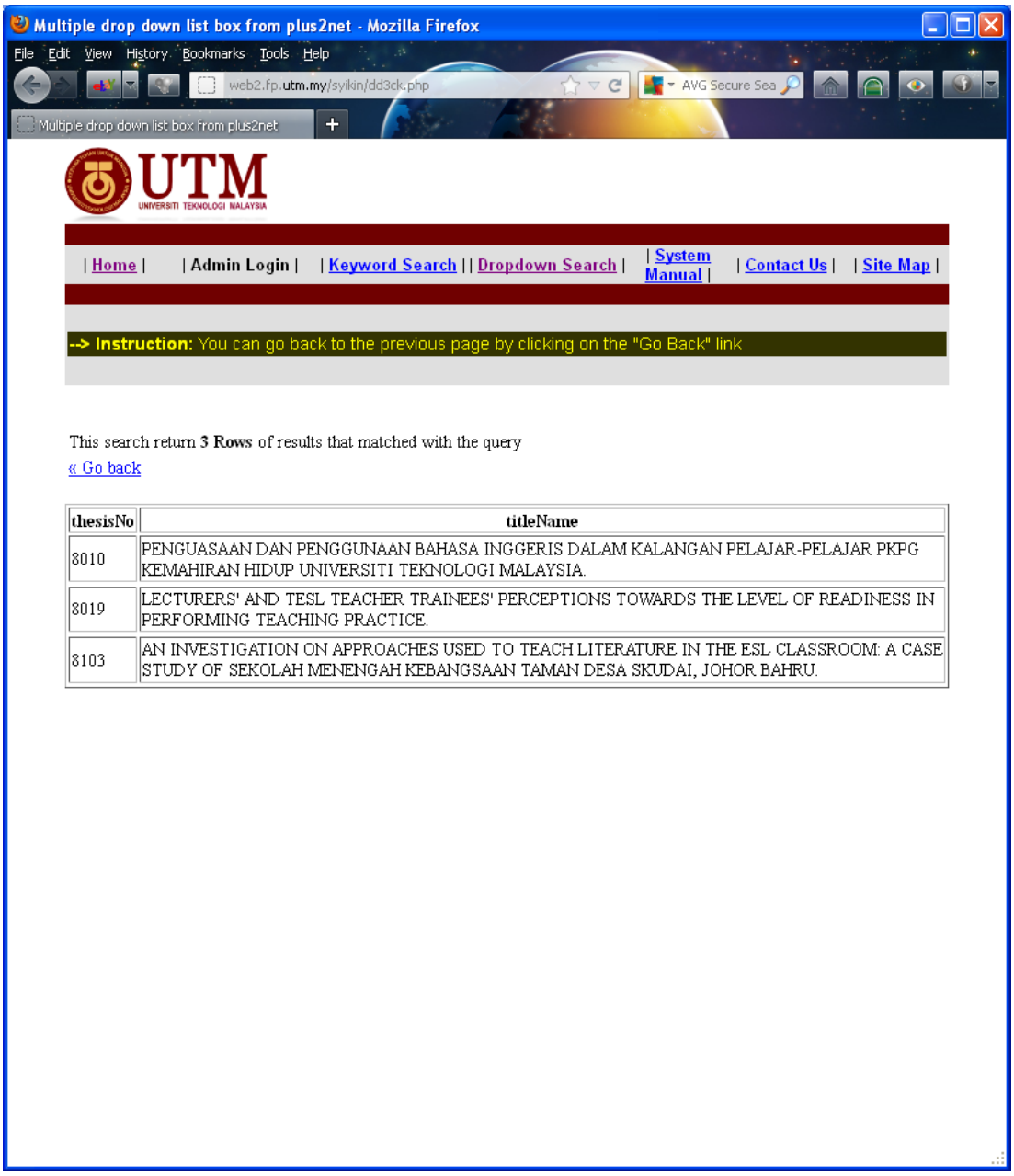


Figure 17: Sample input screen for scenario 3 with the following choice of keywords combination: *Learning → NULL → NULL*

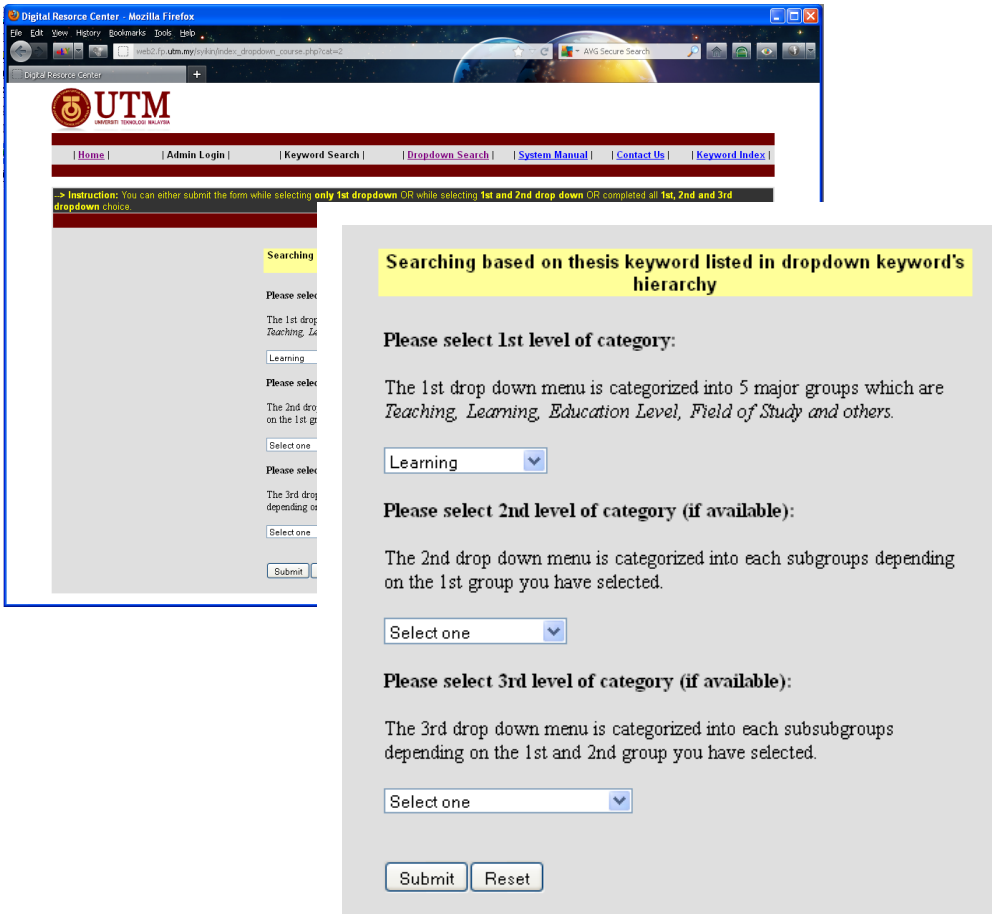
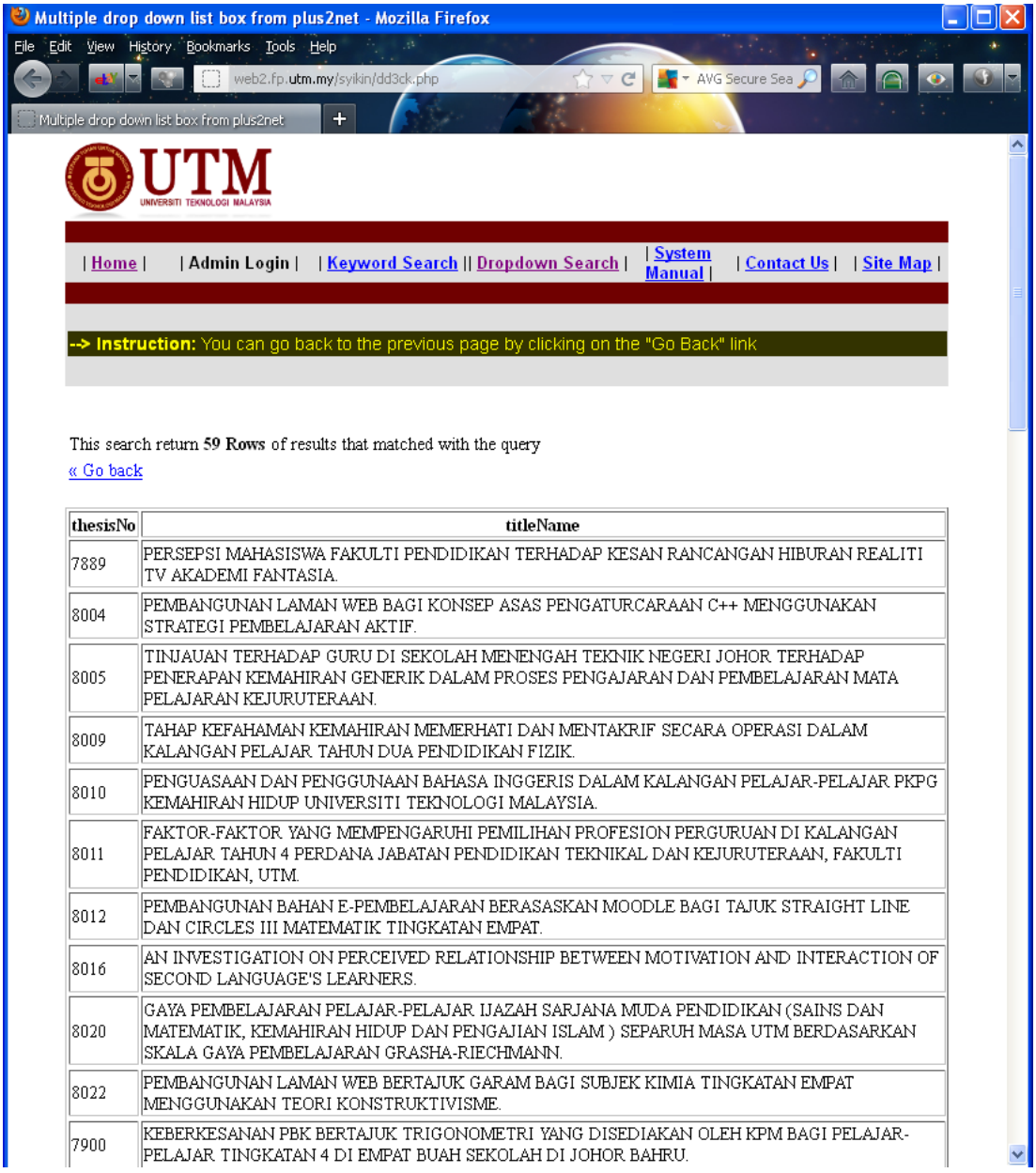


Figure 18: Sample output screen of scenario 3 with the following keywords combination: *Learning → NULL → NULL*



thesisNo	titleName
7889	PERSEPSI MAHASISWA FAKULTI PENDIDIKAN TERHADAP KESAN RANCANGAN HIBURAN REALITI TV AKADEMI FANTASIA.
8004	PEMBANGUNAN LAMAN WEB BAGI KONSEP ASAS PENGATURCARAAN C++ MENGGUNAKAN STRATEGI PEMBELAJARAN AKTIF.
8005	TINJAUAN TERHADAP GURU DI SEKOLAH MENENGAH TEKNIK NEGERI JOHOR TERHADAP PENERAPAN KEMAHIRAN GENERIK DALAM PROSES PENGAJARAN DAN PEMBELAJARAN MATA PELAJARAN KEJURUTERAAN.
8009	TAHAP KEFAHAMAN KEMAHIRAN MEMERHATI DAN MENTAKRIF SECARA OPERASI DALAM KALANGAN PELAJAR TAHUN DUA PENDIDIKAN FIZIK.
8010	PENGUASAAN DAN PENGGUNAAN BAHASA INGGERIS DALAM KALANGAN PELAJAR-PELAJAR PKPG KEMAHIRAN HIDUP UNIVERSITI TEKNOLOGI MALAYSIA.
8011	FAKTOR-FAKTOR YANG MEMPENGARUHI PEMILIHAN PROFESION PERGURUAN DI KALANGAN PELAJAR TAHUN 4 PERDANA JABATAN PENDIDIKAN TEKNIKAL DAN KEJURUTERAAN, FAKULTI PENDIDIKAN, UTM.
8012	PEMBANGUNAN BAHAN E-PEMBELAJARAN BERASASKAN MOODLE BAGI TAJUK STRAIGHT LINE DAN CIRCLES III MATEMATIK TINGKATAN EMPAT.
8016	AN INVESTIGATION ON PERCEIVED RELATIONSHIP BETWEEN MOTIVATION AND INTERACTION OF SECOND LANGUAGE'S LEARNERS.
8020	GAYA PEMBELAJARAN PELAJAR-PELAJAR IJAZAH SARJANA MUDA PENDIDIKAN (SAINS DAN MATEMATIK, KEMAHIRAN HIDUP DAN PENGAJIAN ISLAM) SEPARUH MASA UTM BERDASARKAN SKALA GAYA PEMBELAJARAN GRASHA-RIECHMANN.
8022	PEMBANGUNAN LAMAN WEB BERTAJUK GARAM BAGI SUBJEK KIMIA TINGKATAN EMPAT MENGGUNAKAN TEORI KONSTRUKTIVISME.
7900	KEBERKESANAN PBK BERTAJUK TRIGONOMETRI YANG DISEDIAKAN OLEH KPM BAGI PELAJAR-PELAJAR TINGKATAN 4 DI EMPAT BUAH SEKOLAH DI JOHOR BAHRU.

Figure 19: Sample input screen for scenario 3 with the following choice of keywords combination: *Teaching → Educator → Teacher*

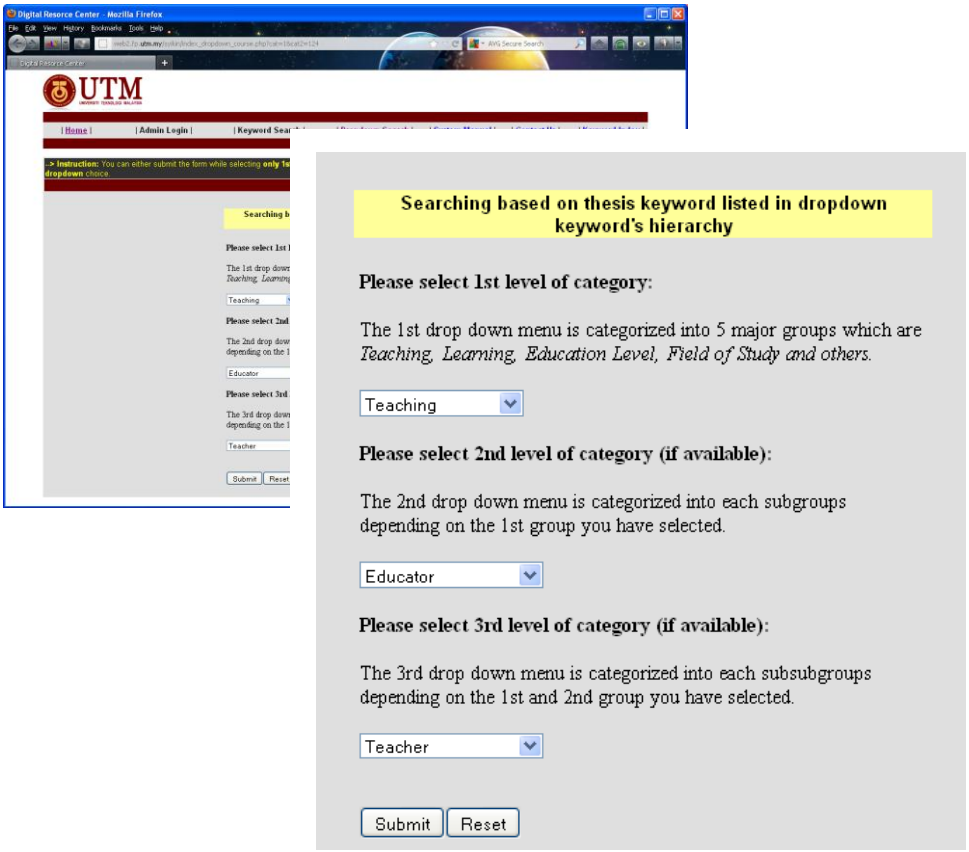
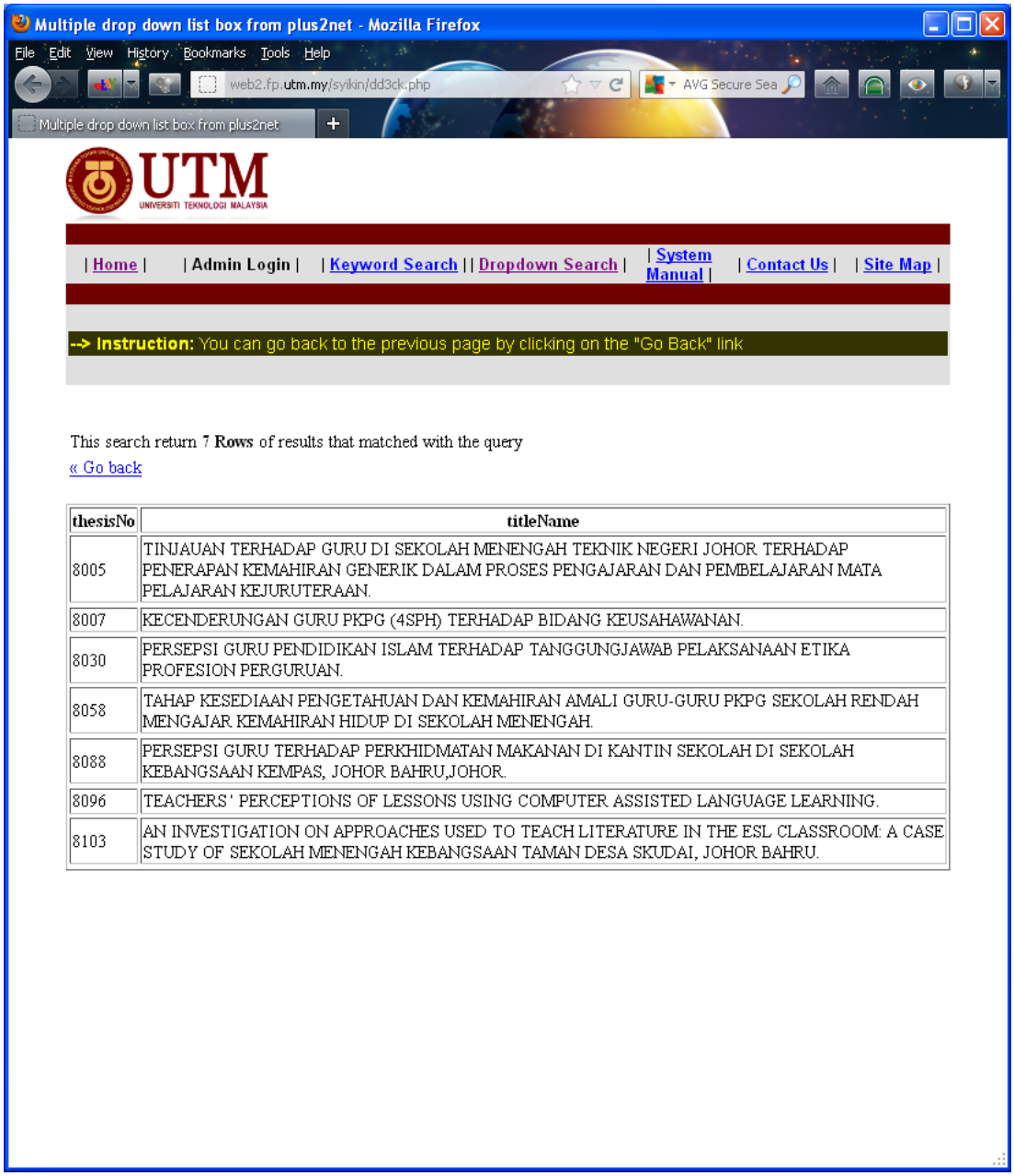


Figure 20: Sample output screen of scenario 3 with the following keywords combination: *Teaching → Educator → Teacher*



thesisNo	titleName
8005	TINJAUAN TERHADAP GURU DI SEKOLAH MENENGAH TEKNIK NEGERI JOHOR TERHADAP PENERAPAN KEMAHIRAN GENERIK DALAM PROSES PENGAJARAN DAN PEMBELAJARAN MATA PELAJARAN KEJURUTERAAN.
8007	KECENDERUNGAN GURU PKPG (4SPH) TERHADAP BIDANG KEUSAHAWANAN.
8030	PERSEPSI GURU PENDIDIKAN ISLAM TERHADAP TANGGUNGJAWAB PELAKSANAAN ETIKA PROFESION PERGURUAN.
8058	TAHAP KESEDIAAN PENGETAHUAN DAN KEMAHIRAN AMALI GURU-GURU PKPG SEKOLAH RENDAH MENGAJAR KEMAHIRAN HIDUP DI SEKOLAH MENENGAH.
8088	PERSEPSI GURU TERHADAP PERKHIDMATAN MAKANAN DI KANTIN SEKOLAH DI SEKOLAH KEBANGSAAN KEMPAS, JOHOR BAHRU,JOHOR.
8096	TEACHERS' PERCEPTIONS OF LESSONS USING COMPUTER ASSISTED LANGUAGE LEARNING.
8103	AN INVESTIGATION ON APPROACHES USED TO TEACH LITERATURE IN THE ESL CLASSROOM: A CASE STUDY OF SEKOLAH MENENGAH KEBANGSAAN TAMAN DESA SKUDAI, JOHOR BAHRU.

APPENDIX E – CODING

index.php

```
1 <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
2 <html xmlns="http://www.w3.org/1999/xhtml">
3   <head>
4     <meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1" />
5     <title>DiRTS System Homepage</title>
6     <style type="text/css">
7       <!--
8       .style4 {font-family: Verdana, Arial, Helvetica, sans-serif; font-weight: bold; font-size: 20px;
        }
9       .style5 {font-family: Verdana, Arial, Helvetica, sans-serif; font-size: 14px; }
10      .style6 {font-family: Verdana, Arial, Helvetica, sans-serif}
11      .style12 {font-family: Verdana, Arial, Helvetica, sans-serif; font-weight: bold; font-size:
        14px; }
12      .style15 {font-size: 1px}
13      .style17 {font-family: Verdana, Arial, Helvetica, sans-serif; font-weight: bold; font-size:
        24px; }
14      .style19 {font-family: Arial, Helvetica, sans-serif; font-size: 14px; }
15      -->
16    </style>
17  </head>
18
19  <body>
20    <table width="90%" border="0" align="center" cellpadding="0" cellspacing="0">
21      <tr>
22        <td colspan="7" bgcolor="#FFFFFF"><img src="image/utmLogo.png" alt="logoutm"
            width="185" height="70" /></td>
23      </tr>
24      <tr>
25        <td height="1" colspan="7" bgcolor="#700000"> </td>
26      </tr>
27      <tr>
28        <td width="99" height="30" bgcolor="#DFDFDF"><div align="center"
            class="style19"><strong>| Home </strong>|</div></td>
29        <td width="162" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style19"><strong>| Admin Login</strong> | </div></td>
30        <td width="154" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style19"><strong>| Keyword Search | </strong></div></td>
31        <td width="163" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style19"><strong>| <a href="index_dropdown.php"
            target="_self">Dropdown Search</a> | </strong></div></td>
32        <td width="100" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style19"><strong>| <a href="about_system.php" target="_self">System
            Manual</a> </strong> | </div></td>
33        <td width="118" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style19"><strong>| <a href="contact_us.php" target="_self">Contact Us</a>
            | </strong></div></td>
34        <td width="109" bgcolor="#DFDFDF"><div align="center" class="style19"><strong>| <a
            href="site_map.php" target="_self">Keyword Index</a> </strong>|</div></td>
35      </tr>
36      <tr>
37        <td height="1" colspan="7" bgcolor="#700000"> </td>
38      </tr>
```

```
39    </table>
40    <p> </p>
41    <table width="90%" border="0" align="center" cellpadding="0" cellspacing="0">
42      <tr>
43        <td height="5" bgcolor="#FFFFFF"><div align="center" class="style4"></div></td>
44        <td height="5" bgcolor="#FFFFFF"><div align="center"><span class="style17">Welcome
              to Digital Resource Thesis Search (DiRTS) System, Faculty Education,  Universiti
              Teknologi Malaysia</span></div></td>
45        <td height="5" bgcolor="#FFFFFF"> </td>
46      </tr>
47      <tr>
48        <td width="11%" height="5" bgcolor="#FFFFFF"> </td>
49        <td width="77%" height="5" bgcolor="#FFFFFF"> </td>
50        <td width="12%" height="5" bgcolor="#FFFFFF"> </td>
51      </tr>
52      <tr>
53        <td rowspan="3"> </td>
54        <td height="119" bordercolor="#999999" bgcolor="#DFDFDF"><div align="center">
55          <p><img src="image/redbook.JPG" alt="book3" width="153" height="103"
                align="absmiddle" /></p>
56        </div></td>
57        <td rowspan="3"> </td>
58      </tr>
59      <tr>
60        <td bgcolor="#DFDFDF"><div align="center"><span class="style12">Fakulti Pendidikan
              Universiti Teknologi Malaysia</span></div></td>
61      </tr>
62      <tr>
63        <td bgcolor="#DFDFDF"><div align="center"><span class="style12">Skudai, Johor,
              MALAYSIA </span></div></td>
64      </tr>
65      <tr>
66        <td> </td>
67        <td bgcolor="#DFDFDF"> </td>
68        <td> </td>
69      </tr>
70      <tr>
71        <td> </td>
72        <td> </td>
73        <td> </td>
74      </tr>
75      <tr>
76        <td> </td>
77        <td><table border="0">
78          <thead>
79            <tr>
80              <th class="style5">DiRTS System offers range of digital resource on Student's
                    Thesis of Faculty of Education, Universiti Teknologi Malaysia .</th>
81            </tr>
82          </thead>
83          <tbody>
84            <tr>
85              <td class="style5"> </td>
86            </tr>
87            <tr>
```

```
88                    <td class="style5"><p align="center" class="style6">You can choose to search
                          the database based on the title of the thesis OR by using the drop down list
                          box features to get the list of thesis details. Click on one of the links
                          below:</p>
89                      <p align="center" class="style6">:: | Keyword Search | <a
                          href="index_dropdown.php" title="Thesis Drop Down List Box Search
                          Option" target="_self">Dropdown Search</a> | :: </p>
90                      <p align="center" class="style6"> </p></td>
91                  </tr>
92                </tbody>
93              </table>    </td>
94            <td> </td>
95          </tr>
96          <tr>
97            <td height="1" bgcolor="#700000"> </td>
98            <td height="1" bgcolor="#700000"><span class="style15"></span></td>
99            <td height="1" bgcolor="#700000"><span class="style15"></span></td>
100         </tr>
101         <tr>
102           <td> </td>
103           <td><div align="center" class="style19">&copy; Copyright by Fakulti Pendidikan,
                      Universiti Teknologi Malaysia</div></td>
104           <td> </td>
105         </tr>
106       </table>
107       <p> </p>
108     </body>
109 </html>
110
111
```

config.php

```
 1 <?php
 2
 3 $dbservertype='mysql';
 4 // hostname or ip of server
 5 $servername='localhost';
 6 // username and password to log onto db server
 7 $dbusername='syikin';
 8 $dbpassword='syikin7';
 9 $dbname='syikin';
10
11 connecttodb($servername,$dbname,$dbusername,$dbpassword);
12 function connecttodb($servername,$dbname,$dbuser,$dbpassword)
13 {
14 global $link;
15 $link=mysql_connect ("$servername","$dbuser","$dbpassword");
16 if(!$link){die("Could not connect to MySQL");}
17 mysql_select_db("$dbname",$link) or die ("could not open db".mysql_error());
18 }
19 ?>
20
```

dd3ck.php

```
<!doctype html public "-//w3c//dtd html 3.2//en">

<html>

<head>
<title>Multiple drop down list box</title>
<meta name="GENERATOR" content="Arachnophilia 4.0">
<meta name="FORMATTER" content="Arachnophilia 4.0">
<style type="text/css">
<!--
.style1 {color: #FFFF00}
.style2 {font-family: Arial, Helvetica, sans-serif}
.style3 {font-family: Arial, Helvetica, sans-serif; font-size: 14px; }
-->
</style>
</head>

<body>
        <table width="90%" border="0" align="center" cellpadding="0" cellspacing="0">
        <tr>
          <td colspan="7" bgcolor="#FFFFFF"><img src="image/utmLogo.png" alt="logoutm"
              width="185" height="70" /></td>
        </tr>
        <tr>
          <td height="1" colspan="7" bgcolor="#700000"> </td>
        </tr>
        <tr>
          <td width="99" height="30" bgcolor="#DFDFDF"><div align="center"
              class="style3"><strong>| <a href="index.php" title="Welcome to Digital Resource
              Thesis Search (DiRTS) System, Faculty Education, Universiti Teknologi Malaysia"
              target="_self">Home</a> </strong>|</div></td>
          <td width="162" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
              class="style3"><strong>| Admin Login | </strong></div></td>
          <td width="154" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
              class="style3"><strong>| <a href="index_keyword.php" target="_self">Keyword
              Search</a> | </strong></div></td>
          <td width="163" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
              class="style3"><strong>| <a href="index_dropdown.php" target="_self">Dropdown
              Search</a> | </strong></div></td>
          <td width="100" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
               class="style3"><strong>| System Manual </strong> | </div></td>
          <td width="118" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
               class="style3"><strong>| <a href="contact_us.php" target="_self">Contact Us</a> |
               </strong></div></td>
          <td width="109" bgcolor="#DFDFDF"><div align="center" class="style3"><strong>| Site
               Map </strong>|</div></td>
        </tr>
        <tr>
          <td height="1" colspan="7" bgcolor="#700000"> </td>
        </tr>
      </table>

      <table width="90%" border="0" align="center" bgcolor="#DFDFDF">
        <tr bordercolor="#000000">
```

```
42            <td> </td>
43
44            <td width="11" height="1"> </td>
45          </tr>
46          <tr>
47            <td height="22" colspan="9" bordercolor="#000000" bgcolor="#333300"><span
                  class="style1 style13"><span class="style16 style2 style1"><strong>--&gt;
                  Instruction: </strong>You can go back to the previous page by clicking on the
                  &quot;Go Back&quot; link </span></span></td>
48        </tr>
49        <tr bordercolor="#FFFFFF" bgcolor="#FFFFFF">
50
51            <td width="13" bgcolor="#DFDFDF"> </td>
52
53
54            <td bordercolor="#000000" bgcolor="#DFDFDF"> </td>
55            <td height="20" bordercolor="#000000" bgcolor="#DFDFDF"></td>
56            <td bgcolor="#DFDFDF"> </td>
57        <td bordercolor="#000000" bgcolor="#DFDFDF" class="style1">        </tr>
58          <tr>
59
60
61
62 <?php
63
64 @$cat=$_POST['subcat1'];
65 @$subcat2=$_POST['subcat2'];
66 @$subcat3=$_POST['subcat3'];
67
68 //echo " cat=$cat <br> subcat2=$subcat2 <br> subcat3= $subcat3 <br> <br>";
69
70
71    mysql_connect('localhost','syikin','syikin7');
72    mysql_select_db('syikin');
73
74    $query= "SELECT DISTINCT idPsm, title
75       FROM syikin.base3
76       WHERE base3.id1=($cat) AND base3.id2=($subcat2) AND base3.id3=($subcat3)";
77
78    $query2= "SELECT DISTINCT idPsm, title
79       FROM syikin.base3
80       WHERE base3.id1=($cat) AND base3.id2=($subcat2)";
81
82    $query3= "SELECT DISTINCT idPsm, title
83       FROM syikin.base3
84       WHERE base3.id1=($cat)";
85
86 $result=mysql_query($query) or $result=mysql_query($query2) or
     $result=mysql_query($query3) or die("Couldn't execute query".mysql_error());
87
88 $num_rows = mysql_num_rows($result);
89
90 echo "<table align='center' width='90%'";
91 echo "<tr><td> This search return <strong> $num_rows Rows </strong> of results that matched
     with the query\n </td></tr>";
92 print "<br>";
```

```
93 echo "<tr><td><a href=\"javascript:history.go(-1)\" title=\"Return to previous page\">&laquo; Go
    back</a></td></tr>";
94 print "<br>";
95 echo "<tr>";
96 echo "</tr>";
97 echo "</table>";
98
99 print "<br>";
100
101 echo "<table align='center' width='90%' border='1'>
102
103 <th>thesisNo</th>
104 <th>titleName</th>";
105
106
107    while($sql_row=mysql_fetch_array($result)) {
108       echo "<center>";
109       echo "<tr>";
110       echo "<td>".$sql_row['idPsm']."</td>";
111       echo "<td>".$sql_row['title']."</td>";
112       echo "<tr>";
113    }
114    echo "</table>";
115
116 ?>
117
118
119
120             </td>
121         </tr>
122       </table>
123
124
125 </body>
126
127 </html>
128
129
```

ajax2.php

```
<?php

@$cat=$_POST['subcat1'];
@$subcat2=$_POST['subcat2'];
@$subcat3=$_POST['subcat3'];

echo " cat=$cat <br> subcat2=$subcat2 <br> subcat3= $subcat3 <br> <br>";

echo "<left>";
echo "Search Result: ";
echo "<br>\n";
echo "<br>\n";


//if (is_numeric($_GET['idpenyelia'])) {

   mysql_connect('localhost','syikin','syikin7');
   mysql_select_db('syikin');

   $query= "SELECT DISTINCT idPsm, title
      FROM syikin.base3
      WHERE base3.id1=$_GET[cat] AND base3.id2=$_GET[subcat2] AND
                  base3.id3=$_GET[subcat3]";

   $query2= "SELECT DISTINCT idPsm, title
      FROM syikin.base3
      WHERE base3.id1=$_GET[cat] AND base3.id2=$_GET[subcat2]";

   $query3= "SELECT DISTINCT idPsm, title
      FROM syikin.base3
      WHERE base3.id1=$_GET[cat]";

$result=mysql_query($query) or $result=mysql_query($query2) or $result=mysql_query($query3)
            or die("Couldn't execute query".mysql_error());


echo "<center>";
   echo "<table align='center' width='90%' border='1'>
<tr>
<th>Id Thesis</th>
<th>Thesis Title</th>
</tr>";

   while($sql_row=mysql_fetch_array($result)) {
      echo "<center>";
      echo "<tr>";
      echo "<td>".$sql_row['idPsm']."</td>";
      echo "<td>".$sql_row['title']."</td>";
      echo "<tr>";
   }
   echo "</table>";
//}
?>
```

index_dropdown.php

```
<html>
  <head>

    <meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
    <link rel="stylesheet" type="text/css" href="style.css">
    <title>DiRTS System Homepage</title>

    <script language="JavaScript" type="text/javascript">
      function display_data(id) {
        xmlhttp=GetXmlHttpObject();
        if (xmlhttp==null) {
          alert ("Your browser does not support AJAX!");
          return;
        }
        var url="ajax.php";
        url=url+"?idpenyelia="+id;
        xmlhttp.onreadystatechange=function() {
          if (xmlhttp.readyState==4 || xmlhttp.readyState=="complete") {
            document.getElementById('result_data').innerHTML=xmlhttp.responseText;
          }
        }
        xmlhttp.open("GET",url,true);
        xmlhttp.send(null);
      }

      function display_data2(id) {
        xmlhttp=GetXmlHttpObject();
        if (xmlhttp==null) {
          alert ("Your browser does not support AJAX!");
          return;
        }
        var url="ajax2.php";
        url=url+"?idpenyelia="+id;
        xmlhttp.onreadystatechange=function() {
          if (xmlhttp.readyState==4 || xmlhttp.readyState=="complete") {
            document.getElementById('div').innerHTML=xmlhttp.responseText;
          }
        }
        xmlhttp.open("GET",url,true);
        xmlhttp.send(null);
      }


      function display_data3(id) {
        xmlhttp=GetXmlHttpObject();
        if (xmlhttp==null) {
          alert ("Your browser does not support AJAX!");
          return;
        }
        var url="ajax3.php";
        url=url+"?idpenyelia="+id;
        xmlhttp.onreadystatechange=function() {
          if (xmlhttp.readyState==4 || xmlhttp.readyState=="complete") {
            document.getElementById('div2').innerHTML=xmlhttp.responseText;
```

```
55              }
56            }
57            xmlhttp.open("GET",url,true);
58            xmlhttp.send(null);
59          }
60
61
62
63          function GetXmlHttpObject() {
64            var xmlhttp=null;
65            try {
66              // Firefox, Opera 8.0+, Safari
67              xmlhttp=new XMLHttpRequest();
68            }
69            catch (e) {
70              // Internet Explorer
71              try {
72                xmlhttp=new ActiveXObject("Msxml2.XMLHTTP");
73              }
74              catch (e) {
75                xmlhttp=new ActiveXObject("Microsoft.XMLHTTP");
76              }
77            }
78            return xmlhttp;
79          }
80      </script>
81      <style type="text/css">
82 <!--
83 .style1 {
84       color: #000000;
85       font-family: Arial, Helvetica, sans-serif;
86       font-size: 16px;
87 }
88 .style4 {font-size: 14}
89 .style12 {font-family: Arial, Helvetica, sans-serif; font-size: 14px; font-weight: bold; color:
   #FFFFFF; }
90 .style13 {
91       color: #660000;
92       font-size: 14px;
93 }
94 .style17 {
95       color: #FF0000;
96       font-weight: bold;
97 }
98 .style18 {font-size: 14; font-family: Arial, Helvetica, sans-serif; }
99 .style20 {font-family: Arial, Helvetica, sans-serif; font-size: 14px; }
100 .style21 {color: #FFFF00}
101 .style22 {
102       font-family: Arial, Helvetica, sans-serif;
103       font-weight: bold;
104       font-size: 16px;
105 }
106 -->
107       </style>
108 </head>
109
```

```
110 <body>
111 <table width="90%" border="0" align="center" cellpadding="0" cellspacing="0">
112     <tr>
113       <td colspan="7" bgcolor="#FFFFFF"><img src="image/utmLogo.png" alt="logoutm"
            width="185" height="70" /></td>
114     </tr>
115     <tr>
116       <td height="1" colspan="7" bgcolor="#700000"> </td>
117     </tr>
118     <tr>
119       <td width="99" height="30" bgcolor="#DFDFDF"><div align="center"
            class="style20"><strong>| <a href="index.php" title="Welcome to Digital Resource
            Thesis Search (DiRTS) System, Faculty Education, Universiti Teknologi Malaysia"
            target="_self">Home</a> </strong>|</div></td>
120       <td width="162" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style20"><strong>| Admin Login | </strong></div></td>
121       <td width="154" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style20"><strong>| Keyword Search | </strong></div></td>
122       <td width="163" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style20"><strong>| Dropdown Search | </strong></div></td>
123       <td width="100" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style19 style20"><strong>| <a href="about_system.php"
            target="_self">System Manual</a> </strong> | </div></td>
124       <td width="118" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style20"><strong>| <a href="contact_us.php" target="_self">Contact Us</a> |
            </strong></div></td>
125       <td width="109" bgcolor="#DFDFDF"><div align="center" class="style20"><strong>| <a
            href="site_map.php" target="_self">Keyword Index</a> </strong>|</div></td>
126     </tr>
127     <tr>
128       <td height="1" colspan="7" bgcolor="#700000"> </td>
129     </tr>
130 </table>
131   <table width="90%" border="0" align="center" bordercolor="#FFFFFF" bgcolor="#DFDFDF">
132     <tr bordercolor="#FFFFFF" bgcolor="#FFFFFF">
133       <td height="22" colspan="7" bordercolor="#CCCC99" bgcolor="#333333"><span
            class="style1 style13"><span class="style21"><strong>--&gt; Instruction:
            </strong>Please read details information and instruction below</span></span></td>
134     </tr>
135     <tr bordercolor="#FFFFFF" bgcolor="#FFFFFF">
136       <td height="22" colspan="7" bordercolor="#CCCC99" bgcolor="#700000"><div
            align="center"><span class="style12">Searching based on thesis keyword listed in
            dropdown keyword's hierarchy </span></div></td>
137     </tr>
138
139     <tr bordercolor="#000000">
140       <td width="5" bordercolor="#DFDFDF"> </td>
141       <td height="21" colspan="5" bordercolor="#DFDFDF" bgcolor="#DFDFDF"> </td>
142       <td width="5" height="21" bordercolor="#DFDFDF"> </td>
143     </tr>
144     <tr>
145       <td width="5" bordercolor="#DFDFDF"> </td>
146       <td height="23" bordercolor="#DFDFDF" bgcolor="#DFDFDF">
147       <div align="center"></td>
148       <td height="20" colspan="3" bordercolor="#D4D0C8" bgcolor="#FFFF99"><div
            align="center" class="style18">
```

```
149          <p>This system is categorized into 5 main categorize which are </p>
150          <p>1.Teaching 2. Learning 3. Education Level 4. Field of Study and 5. others</p>
151          <p> This system is mostly categorized based on <strong>Teaching</strong> and
             <strong>Learning</strong> related keywords identified in the thesis database. Other
             keywords <span class="style17">NOT</span> related to teaching and learning could be
             find under <strong>Others</strong> categories.</p>
152         </div></td>
153         <td height="23" bordercolor="#DFDFDF" bgcolor="#DFDFDF"> </td>
154         <td width="5" bordercolor="#DFDFDF" bgcolor="#DFDFDF"> </td>
155        </tr>
156        <tr>
157
158         <td width="5" bordercolor="#DFDFDF"> </td>
159         <td width="402" height="159" bordercolor="#DFDFDF"
               background="#EFB65B"><p> </p>
160          <p> </p>
161          <p> </p></td>
162         <td width="11" bordercolor="#DFDFDF"> </td>
163         <td width="600" bordercolor="#DFDFDF" bgcolor="#DFDFDF"
               class="style1"><p> </p>
164          <p>Here are some brief description about the <strong>FIVE</strong> main categories.
             </p>
165          <p class="style4"><strong>Teaching</strong></p>
166          <ul class="style4">
167           <li>Anything related to teaching pedagogy, teaching skill, educator, course or subject
                 being teach and etc.</li>
168          </ul>
169          <p class="style4"><strong>Learning</strong></p>
170          <ul class="style4">
171           <li>Could be any matters, person or stuffs that directly related to learning process
                 (Learning Theory, Learning Approach, Learners and etc).</li>
172          </ul>
173         <p class="style4"><strong>Field of Study</strong></p>
174         <ul class="style4">
175          <li>Subject or course undertaken by students or taught by teacher/lecturers are being
                 groups under this categories.</li>
176          </ul>
177         <p class="style4"><strong>Education Level</strong></p>
178         <ul class="style4">
179          <li>Some of the researches are done based on certain level of education for example:
                 research done among secondary technical school or college students.</li>
180          </ul>
181         <p class="style4"><strong>Others</strong></p>
182         <ul>
183          <li class="style4">Most of the keywords which are NOT related to teaching and learning
               terms have been categorized under others. For example teaching practice, athletes,
               academic supervision and etc. </li>
184         </ul></td>
185
186         <td width="9" bordercolor="#DFDFDF"> </td>
187         <td width="407" bordercolor="#DFDFDF"></td>
188         <td width="5" bordercolor="#DFDFDF"> </td>
189        </tr>
190
191
192        <tr bordercolor="#000000">
```

```
193         <td width="5" bordercolor="#DFDFDF"> </td>
194         <td height="53" colspan="5" bordercolor="#DFDFDF"><div align="center"
                class="style22">| <a href="index_dropdown_course.php" target="_self">Proceed to
                begin dropdown search</a> | </div></td>
195         <td width="5" height="53" bordercolor="#DFDFDF"> </td>
196       </tr>
197
198
199
200       <tr>
201         <td width="5" bordercolor="#DFDFDF"> </td>
202         <td bordercolor="#DFDFDF"> </td>
203         <td bordercolor="#DFDFDF"> </td>
204         <td bordercolor="#DFDFDF"> </td>
205
206         <td bordercolor="#DFDFDF"> </td>
207         <td bordercolor="#DFDFDF"> </td>
208         <td width="5" bordercolor="#DFDFDF"> </td>
209       </tr>
210   </table>
211     <p> </p>
212     <p> </p>
213     <p>
214
215       <script language="JavaScript">
216
217         function autoSubmit()
218         {
219            var formObject = document.forms['theForm'];
220            formObject.submit();
221         }
222
223         function autoSubmit2()
224         {
225            var formObject = document.forms['theForm2'];
226            formObject.submit();
227         }
228
229         function autoSubmit3()
230         {
231            var formObject = document.forms['theForm3'];
232            formObject.submit();
233         }
234
235       </script>
236     </p>
237 </body>
238 </html>
239
```

index_dropdown_course.php

```
<?php

// This program is for dropdown menu selection.
// This progrom could be used for specific course selection.
// This progrom could be used for specific supervisor selection.
// This progrom could be used for specific thesis category selection.

require "config.php"; // Database details
?>

<!doctype html public "-//w3c//dtd html 3.2//en">

<html>

<head>
<title>Digital Resorce Center</title>
<meta name="GENERATOR" content="Arachnophilia 4.0">
<meta name="FORMATTER" content="Arachnophilia 4.0">

<SCRIPT language=JavaScript>

function reload(form)
{
var val=form.subcat1.options[form.subcat1.options.selectedIndex].value;
self.location='index_dropdown_course.php?cat=' + val ;
}
function reload2(form)
{
var val=form.subcat1.options[form.subcat1.options.selectedIndex].value;
var val2=form.subcat2.options[form.subcat2.options.selectedIndex].value;

self.location='index_dropdown_course.php?cat=' + val + '&cat2=' + val2 ;
}



function resetForm(){
document.forms[0].search.value = "";
document.forms[0].outlist.value= "";
}



</script>

        <style type="text/css">
            <!--
            .style1 {color: #000000}
.style2 {
        color: #FFFF00;
        font-family: Arial, Helvetica, sans-serif;
        font-size: 14px;
}
.style3 {
        font-family: Arial, Helvetica, sans-serif;
        font-weight: bold;
```

```
57      font-size: 14px;
58 }
59 .style7 {font-family: Arial, Helvetica, sans-serif; font-size: 14px; }
60        -->
61      </style>
62 </head>
63
64   <body>
65      <table width="90%" border="0" align="center" cellpadding="0" cellspacing="0">
66        <tr>
67          <td colspan="7" bgcolor="#FFFFFF"><img src="image/utmLogo.png" alt="logoutm"
              width="185" height="70" /></td>
68        </tr>
69        <tr>
70          <td height="1" colspan="7" bgcolor="#700000"> </td>
71        </tr>
72        <tr>
73          <td width="99" height="30" bgcolor="#DFDFDF"><div align="center"
              class="style7"><strong>| <a href="index.php" title="Welcome to Digital Resource
              Thesis Search (DiRTS) System, Faculty Education, Universiti Teknologi Malaysia"
              target="_self">Home</a> </strong>|</div></td>
74          <td width="162" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
              class="style7"><strong>| Admin Login | </strong></div></td>
75          <td width="154" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
              class="style7"><strong>| Keyword Search | </strong></div></td>
76          <td width="163" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
              class="style7"><strong>| <a href="index_dropdown.php" target="_self">Dropdown
              Search</a> | </strong></div></td>
77          <td width="100" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
              class="style19 style7"><strong>| <a href="about_system.php"
              target="_self">System Manual</a> </strong> | </div></td>
78          <td width="118" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
              class="style7"><strong>| <a href="contact_us.php" target="_self">Contact Us</a> |
              </strong></div></td>
79          <td width="109" bgcolor="#DFDFDF"><div align="center" class="style7"><strong>| <a
              href="site_map.php" target="_self">Keyword Index</a> </strong>|</div></td>
80        </tr>
81        <tr>
82          <td height="1" colspan="7" bgcolor="#700000"> </td>
83        </tr>
84      </table>
85      <p></p>
86      <table width="90%" border="0" align="center" bgcolor="#DFDFDF">
87        <td height="22" colspan="7" bordercolor="#CCCC99" bgcolor="#333333"><span
              class="style1 style13"><span class="style16 style2"><strong>--&gt; Instruction:
              </strong>You can either submit the form while selecting <b>only 1st dropdown</b>
              OR while selecting <b>1st and 2nd drop down</b> OR completed all <b>1st, 2nd and
              3rd dropdown</b> choice.</span></span></td>
88       </tr>
89       <tr bordercolor="#FFFFFF" bgcolor="#FFFFFF">
90        <td height="22" colspan="7" bordercolor="#CCCC99" bgcolor="#700000"><div
              align="center"></div></td>
91
92           <tr bordercolor="#000000">
93          <td> </td>
94          <td height="28" colspan="5"> </td>
```

```
95            <td height="28"> </td>
96          </tr>
97           <tr>
98             <td width="18" bordercolor="#000000" bgcolor="#DFDFDF"> </td>
99            <td width="488" bordercolor="#000000" bgcolor="#DFDFDF"><div
                  align="center"></div></td>
100            <td width="6"> </td>
101            <td width="642" height="33" bgcolor="#FFFF99"><div align="center"
                   class="style3">Searching based on thesis keyword listed in dropdown keyword's
                         hierarchy</div></td>
102            <td width="5"> </td>
103           <td width="465" bgcolor="#DFDFDF"><div align="center"></div></td>
104            <td width="30" bgcolor="#DFDFDF"> </td>
105          </tr>
106          <tr>
107            <td bordercolor="#000000"> </td>
108            <td height="159" bordercolor="#000000"></td>
109            <td> </td>
110            <td bordercolor="#000000" class="style1">
111
112              <table align="left">
113                <tr>
114                  <td>
115
116
117 <?php
118
119
120 ///////// Getting the data from Mysql table for first list box//////////
121 $quer2=mysql_query("SELECT DISTINCT key1,id1 FROM lev123 order by id1");
122 ////////////// End of query for first list box////////////
123
124 /////// for second drop down list we will check if category is selected else we will display all the
            subcategory/////
125 @$cat=$_GET['cat']; // This line is added to take care if your global variable is off
126 if(isset($cat) and strlen($cat) > 0){
127 $quer=mysql_query("SELECT DISTINCT key2,id2 FROM lev123 where id1=$cat order by key2");
128 }else{$quer=mysql_query("SELECT DISTINCT key2,id2 FROM lev123 order by key2"); }
129 ////////// end of query for second subcategory drop down list box ////////////////////////
130
131
132 /////// for Third drop down list we will check if sub category is selected else we will display all
    the subcategory3/////
133 @$cat2=$_GET['cat2']; // This line is added to take care if your global variable is off
134 if(isset($cat2) and strlen($cat) > 0){
135 $quer3=mysql_query("SELECT DISTINCT key3, id3 FROM lev123 where id2=$cat2 order by
    key3");
136 }else{$quer3=mysql_query("SELECT DISTINCT key3 FROM lev123 order by key3"); }
137 ////////// end of query for third subcategory drop down list box ////////////////////////
138
139 print "<br>";
140 print "<b>Please select 1st level of category:</b>";
141 print "<br>";
142 print "<br>";
143 print "The 1st drop down menu is categorized into 5 major groups which are <i> Teaching,
            Learning, Education Level, Field of Study and others.</i>";
```

```
144 print "<br>";
145 print "<br>";
146 echo "<form id=myFormId  method=post name=f1 action='dd3ck.php'>";
147 //////////       Starting of first drop downlist /////////
148 echo "<select name='subcat1' onchange=\"reload(this.form)\"><option value=''>Select
    one</option>";
149 while($noticia2 = mysql_fetch_array($quer2)) {
150 if($noticia2['id1']==@$cat){echo "<option selected
      value='$noticia2[id1]'>$noticia2[key1]</option>"."<BR>";}
151 else{echo  "<option value='$noticia2[id1]'>$noticia2[key1]</option>";}
152 }
153 echo "</select>";
154 //////////////////  This will end the first drop down list ///////////
155
156 print "<br>";
157 print "<br>";
158
159 //////////       Starting of second drop downlist /////////
160 print "<b>Please select 2nd level of category (if available):</b>";
161 print "<br>";
162 print "<br>";
163 print "The 2nd drop down menu is categorized into each subgroups depending on the 1st group
          you have selected.";
164 print "<br>";
165 print "<br>";
166 echo "<select name='subcat2' onchange=\"reload2(this.form)\"><option value=''>Select
         one</option>";
167 while($noticia = mysql_fetch_array($quer)) {
168 if($noticia['id2']==@$cat2){echo "<option selected
      value='$noticia[id2]'>$noticia[key2]</option>"."<BR>";}
169 else{echo  "<option value='$noticia[id2]'>$noticia[key2]</option>";}
170 }
171 echo "</select>";
172 //////////////////  This will end the second drop down list ///////////
173
174 print "<br>";
175 print "<br>";
176
177 //////////       Starting of third drop downlist /////////
178 print "<b>Please select 3rd level of category (if available):</b>";
179 print "<br>";
180 print "<br>";
181 print "The 3rd drop down menu is categorized into each subsubgroups depending on the 1st and
          2nd gr oup you have selected.";
182 print "<br>";
183 print "<br>";
184 echo "<select name='subcat3' > <option value=''>Select one</option>";
185 while($noticia = mysql_fetch_array($quer3)) {
186 echo  "<option value='$noticia[id3]'>$noticia[key3]</option>";
187 }
188 echo "</select>";
189 //////////////////  This will end the third drop down list ///////////
190
191 print "<br>";
192 print "<br>";
193 print "<br>";
```

```
194 echo "<input type=submit value='Submit'>";
195
196 echo "<input type='button' value='Reset' onclick='resetForm();' name='reset' ></form>";
197 ?>
198                         </td>
199                     </tr>
200                 </table>
201             </td>
202         </table>
203
204 </body>
205 </html>
```

C:\3-Current working oct 2011\dec\ThesisFP\details.php

```
1 <html>
 2   <head><title>DiRTS System Homepage</title>
 3   <style type="text/css">
 4 <!--
 5 .style2 {font-family: Arial, Helvetica, sans-serif; font-size: 14px; }
 6 -->
 7   </style>
 8   </head>
 9   <body>
10     <table width="90%" border="0" align="center" cellpadding="0" cellspacing="0">
11       <tr>
12         <td colspan="7" bgcolor="#FFFFFF"><img src="image/utmLogo.png" alt="logoutm"
               width="185" height="70" /></td>
13       </tr>
14       <tr>
15         <td height="1" colspan="7" bgcolor="#700000"> </td>
16       </tr>
17       <tr>
18         <td width="99" height="30" bgcolor="#DFDFDF"><div align="center"
             class="style2"><strong>| <a href="index.php" title="Welcome to Digital Resource
             Thesis Search (DiRTS) System, Faculty Education, Universiti Teknologi Malaysia"
             target="_self">Home</a> </strong>|</div></td>
19         <td width="162" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
             class="style2"><strong>| Admin Login | </strong></div></td>
20         <td width="154" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
             class="style2"><strong>| Keyword Search | </strong></div></td>
21         <td width="163" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
             class="style2"><strong>| <a href="index_dropdown.php" title="Drop Down List Box
             Search Page" target="_self">Dropdown Search</a> | </strong></div></td>
22         <td width="100" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
             class="style2"><strong>| System Manual </strong> | </div></td>
23         <td width="118" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
             class="style2"><strong>| Contact Us | </strong></div></td>
24         <td width="109" bgcolor="#DFDFDF"><div align="center" class="style2"><strong>|
             Keyword Index </strong>|</div></td>
25       </tr>
26       <tr>
27         <td height="1" colspan="7" bgcolor="#700000"> </td>
28       </tr>
29     </table>
30     <p> </p>
31
32     <table width="90%" border="0" align="center">
33       <tr>
34         <td align="center" valign="top"><div align="center"></div></td>
35
36         <?php
37
38         mysql_connect('localhost','syikin','syikin7') or die (mysql_error());
39         mysql_select_db('syikin');
40
41         $line=$_REQUEST['ID'];
42         echo "<table width='70%' border='1' align='center'>";
43
```

```
44          $result=mysql_query("SELECT tb_psm1.ID, AuthorName, title, tahun,
            tb_penyelia_.namaps, tb_kursus_.SingKursus
45                  FROM tb_psm1, tb_penyelia_, tb_kursus_
46                  WHERE (tb_penyelia_.idpenyelia=tb_psm1.idpenyelia OR
                            tb_kursus_.IDkursus=tb_psm1.idkursus) AND tb_psm1.ID='$line'");
47
48          for ($i = 0; $i < 1; ++$i) {
49             $line = mysql_fetch_row($result);
50             echo "<tr><td>ID</td><td>$line[0]</td></tr>";
51             echo "<tr><td>Author Name</td><td>$line[1]</td></tr>";
52             echo "<TR><td>Thesis Title</td><td>$line[2]</td></tr>";
53             echo "<tr><td>Year</td><td>$line[3]</td></tr>";
54             echo "<tr><td>Supervisor Name</td><td>$line[4]</td></tr>";
55             echo "<TR><td>Course</td><td>$line[5]</td></tr>";
56          }
57          echo "</table>";
58
59          ?>
60
61        </tr>
62      </table>
63      <p> </p>
64    </body>
65 </html>
```

about_system.php

```
 1 <html>
 2   <head>
 3
 4     <meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
 5     <link rel="stylesheet" type="text/css" href="style.css">
 6     <title>DiRTS System Homepage</title>
 7
 8     <script language="JavaScript" type="text/javascript">
 9       function display_data(id) {
10         xmlhttp=GetXmlHttpObject();
11         if (xmlhttp==null) {
12           alert ("Your browser does not support AJAX!");
13           return;
14         }
15         var url="ajax.php";
16         url=url+"?idpenyelia="+id;
17         xmlhttp.onreadystatechange=function() {
18           if (xmlhttp.readyState==4 || xmlhttp.readyState=="complete") {
19             document.getElementById('result_data').innerHTML=xmlhttp.responseText;
20           }
21         }
22         xmlhttp.open("GET",url,true);
23         xmlhttp.send(null);
24       }
25       function GetXmlHttpObject() {
26         var xmlhttp=null;
27         try {
28           // Firefox, Opera 8.0+, Safari
29           xmlhttp=new XMLHttpRequest();
30         }
31         catch (e) {
32           // Internet Explorer
33           try {
34             xmlhttp=new ActiveXObject("Msxml2.XMLHTTP");
35           }
36           catch (e) {
37             xmlhttp=new ActiveXObject("Microsoft.XMLHTTP");
38           }
39         }
40         return xmlhttp;
41       }
42     </script>
43     <style type="text/css">
44 <!--
45 .style10 {font-family: Arial, Helvetica, sans-serif; font-size: 14px; }
46 .style15 {font-family: Arial, Helvetica, sans-serif; font-size: 16px; font-weight: bold; }
47 .style17 {font-family: Arial, Helvetica, sans-serif; font-size: 14px; font-style: italic; font-weight:
bold; }
48 .style20 {font-family: Arial, Helvetica, sans-serif}
49 .style21 {font-size: 16px; font-family: Arial, Helvetica, sans-serif;}
50 -->
51     </style>
52 </head>
54   <body>
```

```
55  <table width="90%" border="0" align="center" cellpadding="0" cellspacing="0">
56   <tr>
57    <td colspan="7" bgcolor="#FFFFFF"><img src="image/utmLogo.png" alt="logoutm"
          width="185" height="70" /></td>
58   </tr>
59   <tr>
60    <td height="1" colspan="7" bgcolor="#700000"> </td>
61   </tr>
62   <tr>
63    <td width="99" height="30" bgcolor="#DFDFDF"><div align="center"
        class="style10"><strong>| <a href="index.php" title="Welcome to Digital Resource
        Thesis Search (DiRTS) System, Faculty Education, Universiti Teknologi Malaysia"
        target="_self">Home</a> </strong>|</div></td>
64    <td width="162" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
        class="style10"><strong>| Admin Login | </strong></div></td>
65    <td width="154" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
        class="style10"><strong>| Keyword Search | </strong></div></td>
66    <td width="163" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
        class="style10"><strong>| <a href="index_dropdown.php" target="_self">Dropdown
        Search</a> | </strong></div></td>
67    <td width="100" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
        class="style10"><strong>| System Manual </strong>| </div></td>
68    <td width="118" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
        class="style10"><strong>| <a href="contact_us.php" target="_self">Contact Us</a> |
        </strong></div></td>
69    <td width="109" bgcolor="#DFDFDF"><div align="center" class="style10"><strong>| <a
        href="site_map.php" target="_self">Keyword Index</a> </strong>|</div></td>
70   </tr>
71   <tr>
72    <td height="1" colspan="7" bgcolor="#700000"> </td>
73   </tr>
74  </table>
75  <p> </p>
76  <table width="90%" border="0" align="center" bordercolor="#000000" bgcolor="#FFFFFF">
77       <tr></tr>
78
79       <tr></tr>
80        
81       <tr>
82        <td width="100%" align="left" bgcolor="#DFDFDF">
83
84         <p align="center" class="style10 style20"> </p>
85         <p align="center" class="style15">About System </p>
86         <p align="center" class="style10">This system is develop to assist students with
             minimum search skill to browse the thesis database based on the hierarchy of the
             keyword.<br>
87         The keywords is selected from the thesis's title which have been categorized into main
           areas of Faculty of Education; <br>
88         (1) Teaching and (2) Learning. Others than this two areas, the rest of the keywords can
           be identified under 'others category '. <br>
89         The testing process is going to be done to identify the possibilities of user in determining
           the right keyword based on their logical thinking as well as knowledge evaluation. </p>
90         <p align="center" class="style10">----------
             ----------
             ----------</p>
91         <p align="center" class="style15">Database Category </p>
```

```
92          <p align="center" class="style17">Teaching</p>
93          <p align="center" class="style20">Anything related to teaching pedagogy, teaching skill,
              educator, course or subject being teach and etc.</p>
94                  <p align="center" class="style17">Learning</p>
95          <p align="center" class="style20">Could be any matters, person or stuffs that directly
              related to learning process (Learning Theory, Learning Approach, Learners and
              etc).</p>
96          <p align="center" class="style20"><strong>Field of Study</strong></p>
97          <p align="center" class="style20">Subject or course undertaken by students or taught
              by teacher/lecturers are being groups under this categories.</p>
98          <p align="center" class="style20"><strong>Education Level</strong></p>
99          <p align="center" class="style20">Some of the researches are done based on certain
              level of education for example: research done among secondary technical school or
              college students.</p>
100          <p align="center" class="style20"><strong>Others</strong></p>
101          <p align="center" class="style20">Most of the keywords which are NOT related to
               teaching and learning terms have been categorized under others. For example
               teaching practice, athletes, academic supervision and etc. </p>
102          <p align="center" class="style20">----------
               ----------</p>
103          <p align="center" class="style15">How to use this system </p>
104          <p align="center" class="style17">Dropdown Search </p>
105          <p align="center" class="style10">1. User have to select one of FIVE main categories
               which are Teaching, Learning, Field of Study, Education Level, Others in the 1st level
               of dropdown list.</p>
106          <p align="center" class="style10">2.  User have a choice to select the subcategory of
               the previous level in the 2nd level of dropdown list OR user can proceed with the
               submit button.</p>
107          <p align="center" class="style10">3.  User have a choice to select the subsubcategory
               of the previous level in the 3rd level of dropdown list OR user can proceed with the
               submit button.</p>
108          <p align="center" class="style10">----------
               ----------</p>
109          <p align="center" class="style15">Example</p>
110          <p align="center" class="style21">Assumed that you would like to find for a thesis's
               title written about &amp;quot;learning using computer&amp;quot;. You are going to
               select:</p>
111          <blockquote>
112           <blockquote>
113            <blockquote>
114             <blockquote>
115              <blockquote>
116               <blockquote>
117                <blockquote>
118                 <blockquote>
119                  <blockquote>
120                   <blockquote>
121                    <blockquote>
122                     <blockquote>
123                      <blockquote>
124                       <blockquote>
125                        <blockquote>
126                         <blockquote>
127                          <blockquote>
```

```
128                             <p align="left" class="style21">1st level of dropdown:
                                    <em>Learning</em></p>
129                             <p align="left" class="style21">2nd level of dropdown:
                                    <em>Tool</em></p>
130                             <p align="left" class="style21">3rd Level of dropdown:
                                    <em>Computer  </em></p>
131                           </blockquote>
132                         </blockquote>
133                       </blockquote>
134                     </blockquote>
135                   </blockquote>
136                 </blockquote>
137               </blockquote>
138             </blockquote>
139           </blockquote>
140         </blockquote>
141       </blockquote>
142     </blockquote>
143   </blockquote>
144 </blockquote>
145 </blockquote>
146 </blockquote>
147 </blockquote>
148 <p align="center" class="style21">Once you submit the form, results should be come
      out with some of the titles like below: </p>
149 <table align="center" border="1" width="90%">
150  <tbody>
151   <tr>
152    <th>thesisNo</th>
153    <th>titleName</th>
154   </tr>
155   <tr>
156    <td>7900</td>
157    <td>KEBERKESANAN PBK BERTAJUK  TRIGONOMETRI YANG DISEDIAKAN OLEH KPM
          BAGI PELAJAR-PELAJAR TINGKATAN 4  DI EMPAT BUAH SEKOLAH DI JOHOR
          BAHRU.</td>
158   </tr>
159   <tr></tr>
160   <tr>
161    <td>8031</td>
162    <td>FAKTOR-FAKTOR YANG MEMPENGARUHI PENGGUNAAN KOMPUTER DI
          KALANGAN GURU TEKNIKAL DI SEKOLAH MENENGAH TEKNIK BINTULU
          SARAWAK.</td>
163   </tr>
164   <tr></tr>
165  </tbody>
166 </table>
167 <p align="center" class="style10"> </p>
168 <p align="center" class="style10"> </p></td>
169   </tr>
170  </table>
171  </body>
172 </html>
```

contact_us.php

```
<html>
  <head>

    <meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
    <link rel="stylesheet" type="text/css" href="style.css">
    <title>DiRTS System Homepage</title>

    <script language="JavaScript" type="text/javascript">
      function display_data(id) {
        xmlhttp=GetXmlHttpObject();
        if (xmlhttp==null) {
          alert ("Your browser does not support AJAX!");
          return;
        }
        var url="ajax.php";
        url=url+"?idpenyelia="+id;
        xmlhttp.onreadystatechange=function() {
          if (xmlhttp.readyState==4 || xmlhttp.readyState=="complete") {
            document.getElementById('result_data').innerHTML=xmlhttp.responseText;
          }
        }
        xmlhttp.open("GET",url,true);
        xmlhttp.send(null);
       }
       function GetXmlHttpObject() {
        var xmlhttp=null;
        try {
          // Firefox, Opera 8.0+, Safari
          xmlhttp=new XMLHttpRequest();
        }
        catch (e) {
          // Internet Explorer
          try {
            xmlhttp=new ActiveXObject("Msxml2.XMLHTTP");
          }
          catch (e) {
            xmlhttp=new ActiveXObject("Microsoft.XMLHTTP");
          }
        }
        return xmlhttp;
      }
    </script>
    <style type="text/css">
<!--
.style9 {
    font-family: Arial, Helvetica, sans-serif;
    font-size: 14px;
}
-->
    </style>
</head>

  <body>
  <table width="90%" border="0" align="center" cellpadding="0" cellspacing="0">
```

```
55    <tr>
56     <td colspan="7" bgcolor="#FFFFFF"><img src="image/utmLogo.png" alt="logoutm"
           width="185" height="70" /></td>
57    </tr>
58    <tr>
59     <td height="1" colspan="7" bgcolor="#700000"> </td>
60    </tr>
61    <tr>
62     <td width="99" height="30" bgcolor="#DFDFDF"><div align="center"
           class="style9"><strong>| <a href="index.php" title="Welcome to Digital Resource Thesis
           Search (DiRTS) System, Faculty Education, Universiti Teknologi Malaysia"
           target="_self">Home</a> </strong>|</div></td>
63     <td width="162" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
           class="style9"><strong>| Admin Login | </strong></div></td>
64     <td width="154" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
           class="style9"><strong>| Keyword Search | </strong></div></td>
65     <td width="163" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
           class="style9"><strong>| <a href="index_dropdown.php" target="_self">Dropdown
           Search</a> | </strong></div></td>
66     <td width="100" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
           class="style19 style9"><strong>| <a href="about_system.php" target="_self">System
           Manual</a> </strong> | </div></td>
67     <td width="118" bordercolor="#ECE9D8" bgcolor="#DFDFDF"><div align="center"
            class="style9"><strong>| Contact Us | </strong></div></td>
68     <td width="109" bgcolor="#DFDFDF"><div align="center" class="style9"><strong>| <a
           href="site_map.php" target="_self">Keyword Index</a> </strong>|</div></td>
69    </tr>
70    <tr>
71     <td height="1" colspan="7" bgcolor="#700000"> </td>
72    </tr>
73  </table>
74  <p> </p>
75  <table width="90%" border="0" align="center" bordercolor="#000000" bgcolor="#FFFFFF">
76        <tr></tr>
77
78        <tr></tr>
79         
80        <tr>
81          <td width="100%" align="center" bgcolor="#DFDFDF">
82            <form name="form">
83              <p class="style9"> <strong>If you have any inquiries about this research, you
                    can contact us at:</strong></p>
84              <p class="style9"> <strong>Norasykin Mohd Zaid</strong>
85              <br/>Mobile: +614 1353 1996     Email: nmz056@uowmail.edu.au
86              <br/>      OR
87              <br/> <strong>Sim Kim Lau</strong>
88              <br/>Mobile: +612 4221 4132     Email: simlau@uow.edu.au</p>
89              <p class="style9"><strong>School of Information Systems and Technology
                    (SISAT)</strong>
90              <br/>University of Wollongong
91              <br/>Wollongong NSW 2522 Australia </p>
92
93              <p>    </p>
94            </form>            </td>
95        </tr>
96  </table>
```

```
97    </body>
98 </html>
```