FACTORS AFFECTING QUERY FORMULATION IN WEB INFORMATION SEARCH: A CASE STUDY OF THE UNIVERSITY OF MORATUWA

A D B KUMARA

(2009/MISM/19)

Thesis submitted in partial fulfilment of the requirements for the degree of Master of Information Systems Management

Faculty of graduate studies
University of Colombo
Colombo

NOVEMBER 2011

DECLARATION

I certify that this Dissertation does not incorporate without acknowledgement any material previously submitted for the Degree or Diploma in any University, and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text

Date: .....

A D B Kumara

The undersigned, have supervised the dissertation entitled **FACTORS AFFECTING QUERY FORMULATION IN WEB INFORMATION SEARCH: A CASE STUDY OF THE UNIVERSITY OF MORATUWA** presented by A D B Kumara, a candidate for the degree of Masters in Information Systems Management, and hereby certify that, in my opinion, it is worthy of submission for examination.

Date: ....

Dr. W Seneviratne

Supervisor

Abstract

The World Wide Web, which is exponentially growing daily, is also known to be the richest source of information. The Web is universally adopted by university academic communities as a part of their information seeking motivated by a wide variety of needs including academic, professional, personal or entertainment etc. The ability to search and retrieve information from the Web effectively and efficiently is a challenge as current search tools retrieve too many documents of which only a small fraction is relevant to the user query. Studies have revealed that well formulated queries are one of best solutions for this challenge. This study is an attempt to find out the factors affecting query formulation on Web information searching of university academic community. The study is compiled with data gathered from questionnaires of 255 undergraduate students, 107 postgraduate students, 40 NDT students and 30 teaching faculty members. The total response rate is 98.63%. The results indicate Web experience, topic familiarity and search strategies as the identified factors affecting query formulation. The average number of stop words per query decreases when Web used experience increases and the time spent to formulate a query increases when the Web used experience increases. Also, the time spent per query decreases with a familiar topic and he/she desire to use synonyms. Recommendations were made to conduct training programs on available Web information sources/resources and Web searching strategies.

Acknowledgement

It is with deep appreciation that I acknowledge the assistance that was extended to me in carrying out this research project and I wish to pay my tribute for all of them for their tremendous services.

May I in the first instance express my deepest gratitude and sincerest thanks to my supervisor, Dr. W. Seneviratne Librarian, Open University of Sri Lanka for her invaluable guidance and encouragement through this project.

I wish to gratefully acknowledge, Dr. C. Jayasundara, the Course Coordinator and Senior Assistant Librarian, University of Colombo for his continuous help and guidance in all endeavours and all the other lecturers contributed through their opinions and helped in very many ways.

I wish to specially acknowledge, Mrs. S. C. Jayasuriya, former Librarian, University of Colombo who gave me a great assistance to successfully compile this research.

I am greatly indebted to Mrs. R.C. Kodikara, Librarian, University of Moratuwa for her ready assistance and encouragement and care throughout this study. It was a pleasant experience to take part in discussions with Dr. R. Gamage and Ms. T. Senewirathne Senior Assistant Librarians University of Moratuwa, the academic and administrative staff of the Library, University of Moratuwa. Their constructive comments helped me to develop this study.

I like to appreciate the help and cooperation rendered by all the Heads of the academic Departments/Divisions of University of Moratuwa and Institute of Technology. I wish to place a record on the generous assistance accorded to me in providing statistical data by the Heads and the staff of Academic Establishment Division, Examination Division, Institute of Technology, University of Moratuwa.

My special thanks go to Ms. G.D.M.N. Samaradiwakara and Ms. R. Samarawickrama Senior Assistant Librarians University of Sri Jayawardenapura for helping me in numerous ways throughout the project.

Finally I wish to extend my thanks to my wife, daughter and son for their encouragement, patience and understanding support given during this study.

Table of Contents

List of Figures

List of Tables

Chapter 1 - Introduction

World Wide Web (Web or WWW) has facilitated extensive and timely dissemination of information and turned the world into a global village. The university academic community plays a key role in accessing information for all their academic and daily activities in this global village.

The increase in information available on the Web has a huge impact on access to information and on information seeking behavior of the university academic community. One of the major tools for web information access is the search engine. Searcher formulates a query on the search engine and abundant information is visible on the computer screen. Most of it is not relevant, reliable or up to date. Tracing such a lot pages of pages is a time consuming task for lecturers as well as students with a busy schedule in the university. Therefore, the ability to search and retrieve information from the Web effectively and efficiently is an enabling issue that requires full potential.

Considerable research work has been done recently worldwide to find out ways and means to solve this problem of Web information retrieval. One of the key aspects of improving the efficiency and effectiveness of Web search involves well formulated queries. Thus it is very vital to consider the factors affecting query formulation in Web information search in order to design better search tools and ways of retrieval and access.

1.1 University academic community and the information seeking behavior

Universities are places where the learning and research community always deal with valuable, as well as current information. This community concentrates on an enormous amount of information that is collected, analyzed and transformed into valuable knowledge (Bhatti, 2010). Schneiderman, Byrd and Croft (1997) defined 'information needs' as the perceived need for information that leads to someone using an information retrieval system to obtain output expected by them. Information seeking behavior includes personal information needs as well.

Information seeking behavior is the purposive seeking for information as a consequence of a need to satisfy some goal (Patitungkho and Deshpande, 2005). In the course of seeking, the individual may interact with manual information systems (such as newspaper or a library), or with computer-based systems (such as the Web) (Wilson, 2000). Information seeking behavior is expressed in various forms, from printed material to research and experimentation. Further, it is described as the process or activity of attempting to obtain information in both human and technological contexts.

University academic community makes active and intentional attempts to retrieve up-to-date information from the various media available in the library such as books, encyclopedias, journals and more currently electronic media. In acquiring knowledge, the user needs to associate with an effective information retrieval system, which is inherently predicated on users searching for information. The first systematic solution to the problem of finding the desired information from a large information collection was developed about 4,000 years ago by librarians who kept track of "books" by cataloging them by author and title (Cleveland & Cleveland, 1990). Searching through the catalog to

find a book was a marked improvement from the physical search of actual books, but it required the searcher to know not only what book s/he was looking for, but also know it by its author and title.

The next solution came in the 16th century in the form of crude indexes, which were a list of keywords (e.g., topic, subject) with pointers to documents (Wheatley, 2002). The basic idea was that people could find the desired information by selecting the “appropriate" keyword entry in the index, which would list the documents related to it. Though this keyword index strategy expanded the capability of the catalog method by allowing the searcher to find a set of items related for a given concept (i.e., keyword), it introduced the problem of ambiguity in representation. Since there was no explicit rule for assigning keywords to documents, the choice of keywords for a given document depended heavily on the subjective word choice for a particular interpretation of the document. The obvious problem with this approach was the difficulty of selecting the "appropriate" keywords to express the information need.

1.2 Information retrieval in new technological arena

With the advent of computers in the 1950s came computerized information searching, where the strategy of using the terms in the document collection to create the index was explored (Luhn, 1957). Over the past ten years, the Internet has received growing importance all over the world. It gives access to an endless amount of information. Just as the Web has grown rapidly, the searching on the Web for information access has become an efficient tool; not only for information access, but web functionalities spreading across difference aspects of lives of people as well. With the growth of WWW, a need of

complexity raised in the information retrieval criteria/strategy Web context, the human-computer interaction factors and the cognitive aspects play a significant role.

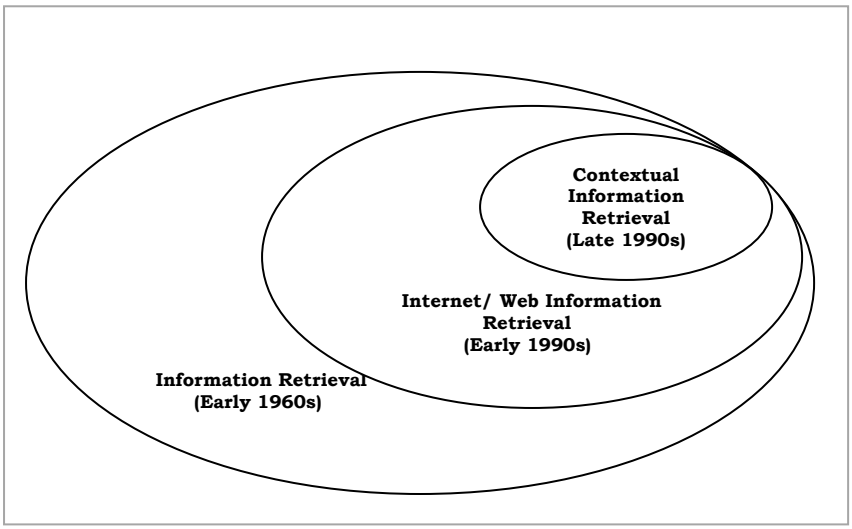


Figure 1.2.1: Evolution of information retrieval (Yates & Neto, 1999)

Figure 1.2.1 illustrates the evolution of information retrieval, where Internet/ Web information retrieval is a sub-discipline within the general information retrieval arena. As Yates & Neto stated in 1999, Internet/ Web information retrieval uses either search engines which index a portion of documents as a full-text databases, or web directories, which classify selected Web documents by subject to facilitate the identification of relevant information. Due to the exponential growth of information on the Web, the search engines had to face a huge challenge. In the late 1990s, contextual information retrieval was introduced to address these challenges. Contextual information retrieval is the gaining of appropriate information to fulfill the information needs exactly. This has been made easier with the expansion of the World Wide Web and the advent of modern and inexpensive graphical user interfaces and mass storage devices. Therefore, the information retrieval or resource discovery in web context is a long-term, multi-faceted, iterative process with complex and dynamic requirements of modern users (Lagoze, 1997).

The new digital environment not only forces people to apply more than one type of information-seeking strategies, but also requires people to change from one information seeking strategy to another in the information seeking process. Therefore, searching for information on the Web is an important aspect which is very different from that undertaken in more traditional retrieval environments (Spink et al., 2001).

1.3 Web information searching and Query formulation

Information searching is a complex process of articulating an information need, often ambiguous, into precise words and relationships that match the structure of the system being searched, bridging the knowledge gap that exists within an information need of users at a given time. It consists of problem identification, need articulation, query formulation, and results evaluation (Sutcliffe, 1998).

Searching is the most important application among all web-based applications. However, satisfying the information needs through the Web is not always an easy and straightforward process. It is very time consuming and may be an exhausting experience for a novice user who may culminate unsatisfactory results as well.

According to Borgman (1987), the knowledge required for Web searching has been divided into two components; mechanical and conceptual. Mechanical aspects of searching are the syntax for entering search terms, the implementation of various functions, and methods of negotiating and altering search statements. Conceptual aspects of searching or the “how and why” are; when to use which function, alternative search paths etc. It involves the cognitive process of a user and problem solving behavior of an individual who involved in Web searching.

Basically, three levels, conceptual, linguistic and string level can be differentiated in query formulation (UMLS, 1994). A concept may be represented by several expressions and those expressions may be varied according to the search environment, language. Each concept in a query is represented by one or more natural language expressions at the linguistic level. At string level each expression is represented by one or more search strings. Therefore one of major issues in varying queries on Web is language or the linguistic level of the searcher.

Usage of Web searching spread across heterogeneous populations:

- Some only focus on entertainment, while others require accurate and efficient search facilities for professional purposes.
- There are computer novices, intermediate users, as well as highly-skilled experts.

Although there are different methods for information retrieval through the Web environment, the most common way is searching through different web-based search facilities; Search Engines. There are approximately 3,200 search engines on the Web with a handful dominating in terms of usage (Sullivan, 2009). Most search engines are designed to serve the heterogeneous population as a whole. They offer amenities for advanced query formulation such as Boolean operators, phrase searching and Ad hoc queries as well.

However, current information retrieval systems such as Web search engines (e.g., Google), online databases (e.g., ScienceDirect), online public access catalogues (OPAC) and digital libraries (e.g., American Memory from Library of Congress) are designed mainly to support query formulation and limited browsing. This process is affected by

environmental (e.g., the database and the search topic), searcher (e.g., online search experience), search process (e.g., commands used), and search outcome variables (e.g., precision and recall) (Fenichel, 1981).

1.4 Significance of the study

It is clear that information searching and retrieval of expected results is a complex activity. The university academic community is dealing with higher education that comprises of teaching, learning and research. Teaching faculty, as well as students, is seeking for information for their academic activities and day to day purposes. They hardly find time for seeking information from non-Web sources. On the other hand, Web supplies plenty of various kinds of information in a very short time. Therefore, the academic community is compelled to find information through the Web. Their ability to proficiently searching the Web is very useful in their information seeking career to save time. Query formulation efficiency is identified as a key concern in Web information searching.

Although interactive query formulation has been actively studied in the laboratory, little is known about the actual behavior of Web searchers on query formulation. Search engine designers, and information skills trainers do not have sufficient knowledge on possible practical problems engaged by different user groups. Therefore, a need lies to study and explore the factors affecting query formulation in relation to different user categories in the Sri Lankan context, since no such a study has been done before.

Hence, librarians have the responsibility to introduce training programs which will help them to increase the efficiency of web searching. This study can identify useful criteria to

follow and aspects to focus on, in teaching information searching by focusing on the variables affecting query formulation of the searcher.

1.5 Statement of the problem

The galloping growth of the Web leads to build up a huge and complex repository of information for diverse communities and for different provisions of them as well. The university academic community is one of the front liners of this multidisciplinary web user population that consists of experienced and novice users. They construct searches differently. The required information may exist on the Web, but due to various reasons the user may not be able to find it. Exploring possible reasons for this failure (locating information through a query formulation) is a challenging issue. Non identification of these failures and reasons for the failures of University of Moratuwa academic community would lead to losing the leaving time unnecessary. It affects their ability to achieve academic targets. Therefore this study can be used to rectify their problems and thus leading them for better use of information, and for better achievements.

1.6 Theoretical background

The field of information retrieval has developed as two largely unconnected but related subfields: one that focuses on systems aspects, and the other that focuses on the human, cognitive, and interactive aspects. The field of human information behaviour is related to the cognitive approach to interactive information retrieval and seeks to investigate the broader issues related to human processes for seeking and using information. However, interactive information retrieval research and human information behaviour research have been largely unconnected despite their mutual interest in areas of human information-related behavior. The appropriate integration of elements of both fields is growing in importance, particularly to further the development of more effective, theoretical models,

and Web, and IR systems design and evaluation (Spink, 1999; Vakkari, 1999; Wilson, 1999). But Application of information seeking theories and models has also been limited in interactive information retrieval (Joho, 2009).

Interactive search episodes are represented by aspects from various interactive information retrieval models, including models by Ingwersen (1992, 1996), Belkin, Cool, Stein, and Theil (1995), and Saracevic (1996, 1997), and in studies by such researchers as Bates (1989) and Fidel and Soergel (1983). Over time, movements or shifts may take place during interactive search episodes and between searches, including changes in tactics, the definition of the information problem, strategies, terms, feedback, goal states, or uncertainty. Search episodes provide the informational framework to the problem solving process through which the user's uncertainty level is reduced.

The cognitive model of Ingwersen

The information-seeking and retrieval research framework has been developed for over a decade by Ingwersen (1992, 1996, 1999) and Ingwersen and Järvelin (2005). Ingwersen (1992, 1996) developed and enhanced the cognitive model of information retrieval interaction, which set up the foundation for the integrated information systems and retrieval research framework proposed by Ingwersen and Järvelin (2005). The five components (an individual user's cognitive space, a user's social-organizational environment, the interface/intermediary, the information objects, and the information retrieval system setting), the cognitive transformation and influence from one component to another, and the interactive communication of cognitive structures via an interface or intermediary constitute the cognitive model of interaction. In this model, "cognitive structures are manifestations of human cognition, reflection or ideas. In information

retrieval they take the form of transformation generated by a variety of human actors" (Ingwersen, 1996, p. 8).

The cognitive approach to information retrieval theory makes use of the major discoveries, experiences, investigations, and experimental tests made over more than half a century in information science. A cognitive effect is a form of change in our cognitive state or process such as perception, learning, problem solving, memory, attention, and language caused by an event. However, the cognitive model of Ingwersen has serious consequences for the understanding of information retrieval and the research in the field.

The cognitive framework for information searching

Process model consists of two components, activity and strategies. Information searching activity and strategies consists of problem identification, articulating needs, query formulation and evaluating results. Following model summarized the cycle of information seeking and its principal activities.

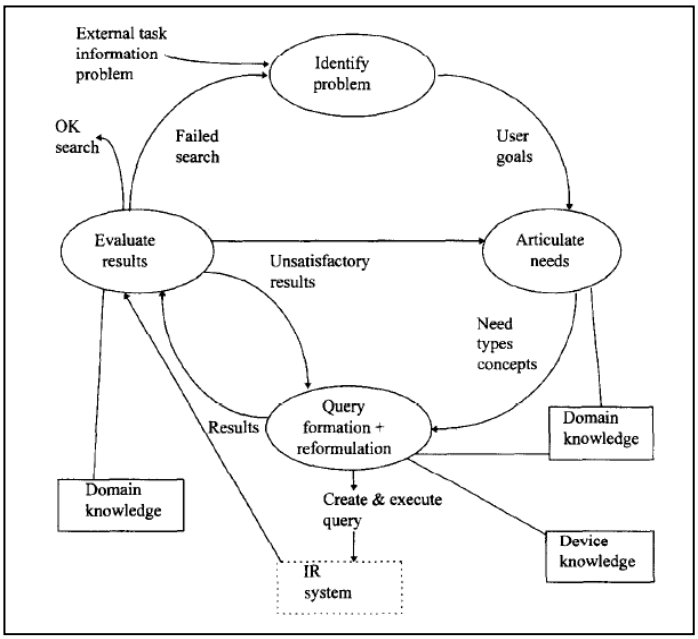


Figure 1.6.1: Process model of information searching activities and knowledge sources (Sutcliffe and Ennis,1998)

Problem identification: involves identifying the initial goal or information need. If the problem or information need is complex, decomposition methods are used to split the problem into smaller components.

Articulating needs: once a need has been identified it may be expressed as concepts or high level semantics. These are refined into lower level terms which are utilized in queries.

Query formulation: the complexity of this activity depends on the sophistication of the information retrieval system and the user's skill in generating queries. Hence complex queries can be formed if the user is skilled in Boolean query languages and such a language is available in the retrieval systems. Alternatively, query formulation may be trivial with a hypertext system when the user merely recognizes search terms as hotspots and follows information links. Two sub-activities may be recognized; first identification of search terms and then transforming these into the query language supported by the search system.

Evaluating results: once information is retrieved the user has to decide whether to accept the retrieved results or to continue searching. Three sub-tasks are involved. First, the user has to scan the result set or examine the contents in detail. Various scanning and sampling strategies are possible. Then a decision has to be reached as to how useful the retrieved results are and whether they are sufficient to meet the need. Once this decision is reached the user has to elect to either accept the results or decide how the query should be changed. Evaluation activity will depend on the richness of the user's domain knowledge; the more you know about a subject area, the better one can determine the relevance of the search results. However, the complexity of this activity is also determined by characteristics of the machine.

General model for information searching process (Spink et al. 2002)

Table 1.6.1: General model information seeking and searching process (Spink et al., 2002)

Event	Variables
Information seeker has an information problem to resolve	Information seeker pre-search characteristics Cognitive style Problem statement Knowledge level Problem-solving stage Information seeking stage Uncertainty level
Information seeking process related to information problem	Information seeking behaviors Information seeking models
Information seeker formulates their information problem into a question	Question statement Question analysis
Pre-search interaction with a search intermediary	Intermediary characteristics
Formulation of the search strategy (terms and tactics)	Pre-search characteristics: information seeker
Search activity and interactions	Search strategy Search characteristics Search processes Successive Searchers
Delivery of responses to the information seeker	Items retrieved Forms delivered
Evaluation of output	Relevance Utility
Information seeker evaluation of impact of search	Information seeker post-search characteristics Problem statement Knowledge level Problem-solving stage Information seeking stage Uncertainty level

The general model adopted for the study is a basis for the development towards integration of interactive information retrieval within information seeking contexts, and for exploring information seekers' interactive search episodes within their changing information- searching behavior. One may advance the points that are the most central for deeper understanding of information retrieval and the study of the phenomena associated with information retrieval.

1.7 Conceptual framework of the study

According to the literature reviewed, the process of information searching in Web which is depicted in the following model consists of problem identification, need articulation, query formulation and results evaluation (Sutcliffe, 1998; Hwang, 2011) and they are mainly affected by factors; environmental, searcher, search process, and the search outcome (Fenichel, 1981) The model is given in figure 1.7.1.

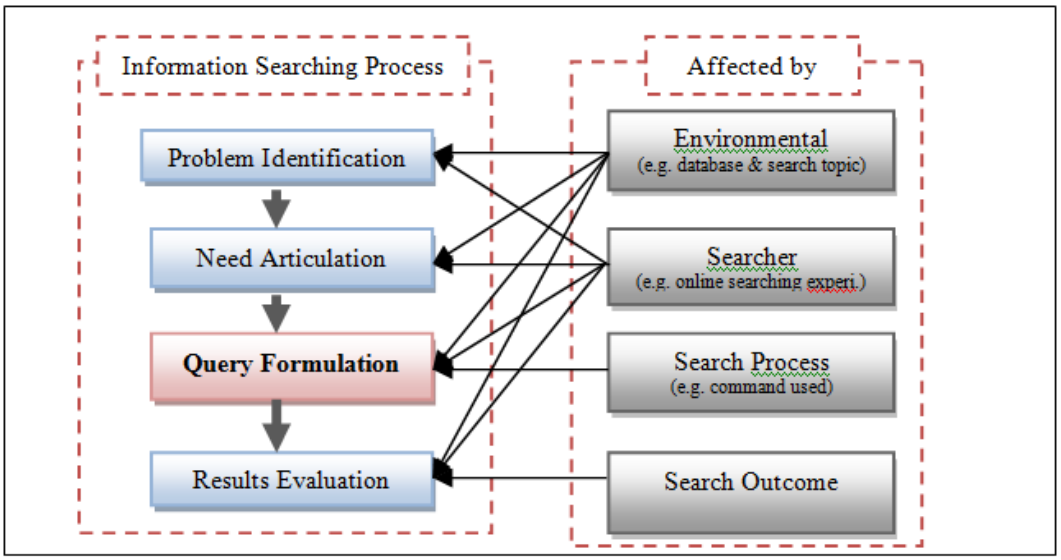


Figure 1.7.1: Process of information searching on Web

Since the current study is focused on investigating the factors affecting **query formulation** in web searching among the university academic community, it was further explored using prior studies.

Based on Aula (2003) who studied query formulation, three main factors are expected to have a significant impact. These are; media expertise, domain expertise, and the type of search task (Ingwersen, 2010).

Media (Web) expertise: Familiarity with the search environment, search engine expertise, computer expertise, expertise in information retrieval are factors that are

expected to have an effect on query formulation. Under this criterion, the number of search terms used and the type of query (broad or precise) the user prefers were considered as measures. Furthermore, the search style; straight to information and navigating to information were considered as measures.

Domain expertise: Under this criterion the usage of synonyms, more accurate terms and a higher number of terms in search queries were considered as measures.

Type of search tasks: Search tasks can be divided into three broad categories; fact-finding, exploratory, and comprehensive search tasks. Using precise terms or phrases in the query, and formulating a query consisting of several terms, broader terms, and term truncation, or other methods for maximizing the recall, and using advanced operators (e.g., Boolean OR and AND) were the factors considered.

Considering all the factors found in literature traced, the researcher could develop a model framework for the study as illustrated in Figure 1.7.2:

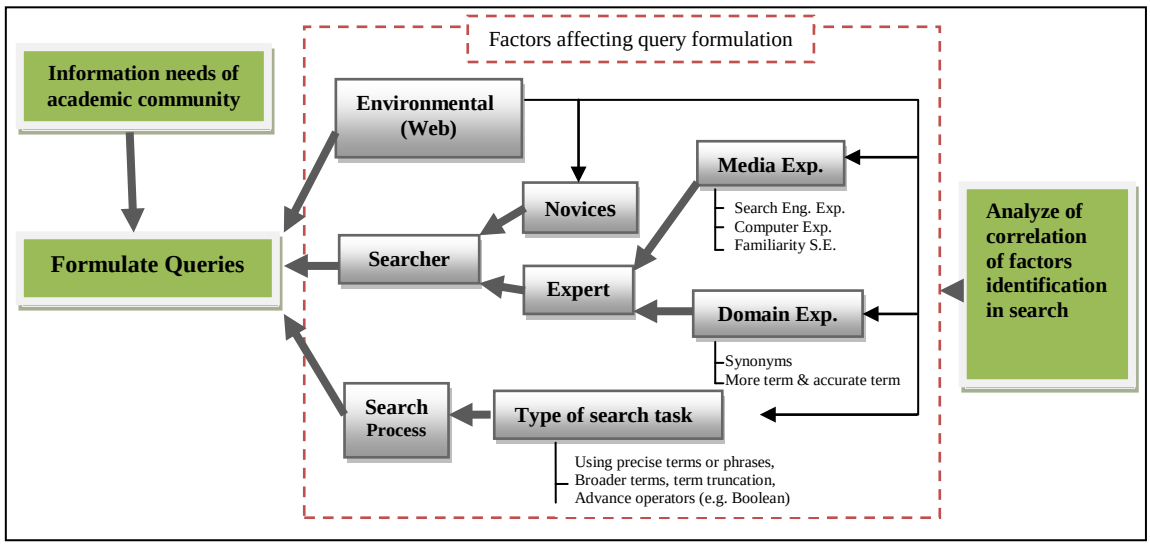


Figure 1.7.2: The Model framework for the study

1.8 Objectives of the study

1.8.1 General objective

To investigate formulation of queries in Web searching by the academic community of the University of Moratuwa and factors affecting the formulation of queries.

1.8.2 Specific objectives

The specific objectives of the study are:

1 To identify the information purpose for searching the web by the academic community in the University of Moratuwa.

2 To investigate the factors affecting using queries by the academic community in the University of Moratuwa.

3 To examine how the academic community conceptualize search strategy.

4 To identify the types of web information search used by the academic community in the University of Moratuwa.

1.8.3 Research Questions

To enable the investigation five research questions were formulated:

RQ1. What are the purposes of searching web information by the academic community in the University of Moratuwa?

RQ2. What are the factors affecting query formulation in searching information on web?

RQ3. How does the academic community formulate queries in web searching?

RQ4. How do they judge relevancy of information?

RQ5. What types of web information are retrieved by the academic community in the University of Moratuwa?

References

Aula, A. (2003) Query formulation in web information search. *Proceedings of the IADIS international conference WWW/Internet 2003*, Algarve, Portugal, November 2003. IADIS Press, Lisbon, Portugal, p. 403–410

Bates, M.J. (1989) The design of browsing and berry picking techniques for the on-line search interface. *Online Review*, **13(**5), p.407–431.

Belkin, N.J. et al. (1995) Cases, scripts and information seeking strategies: on the design of interactive information retrieval systems. *Expert Systems with Applications*, 9, p.379–395.

Bhatti, Rubina. (2010) Information needs and information- seeking behavior of faculty members at the Islamia University of Bahawalpur. Available from: http://www.webpages.uidaho.edu/~mbolin/bhatti.htm [Accessed 30th Dec 2010].

Borgman, Christine L. (1987) The study of user behaviour on information retrieval systems. *SIGCUE* [online], 19(3-4) Springer/Summer 1987. p.35-48. Available from: http://delivery.acm.org/10.1145/970000/964086/p35-borgman.pdf?key1=964086&key2=8272247921&coll=DL&dl=ACM&CFID=9591540&CFTOKEN=53892468 [Accessed 4th March 2010].

Cleveland, D. B and Cleveland, A.D. (eds) (1990) Introduction to Indexing and Abstracting. Englewood, CO: Libraries Unlimited, Inc.

Fenichel, C. H. (1981) Online searching: Measures that discriminate among users with different types of experiences. *Journal of the American Society for Information Science*, 32(1), p.23-32.

Fidel, R. and Soergel, D. (1983) Factors affecting online bibliographic retrieval: A conceptual framework for research. *Journal of the American Society for Information Science*,34(3), p.163-180.

Hwang, G.-J. and Kuo, F.-R. (2011) An information-summarizing instruction strategy for improving the web-based problem solving abilities of students. *Australasian Journal of Educational Technology*, 27(2), p.290-306.

Ingwersen, P. (1992) Information retrieval interaction. London: Taylor Graham.

Ingwersen, P. (1996) Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation*, 52(1), p.3-50.

Ingwersen, P. (1999) Cognitive Information Retrieval. *Annual Review of Information Science and Technology (ARIST)*, 34, p.3-52.

Ingwersen, P. et al. (2010) Panel: Perspectives on adaptivity in information retrieval interaction (PAIRI). *Proceedings of the 73rd ASIS&T Annual Meeting*, 47, p.1-5.

Ingwersen, P. and Järvelin, K. (2005) Information Retrieval in Context – IRiX. *ACM SIGIR Forum*, 39(2), p.31-39.

Joho, H. (2009) Cognitive effects in information seeking and retrieval. *ECIR 2009 Workshop on Contextual Information Access, Seeking and Retrieval Evaluation*, Available at: http://www.dcs.gla.ac.uk/~hideo/pub/cirse09/cirse09.pdf [Accessed 20th September 2010]

Lagoze, C. (1997) From static to dynamic surrogates: resource discovery in the digital age. *D-Lib Magazine* [Online], Available: http://www.dlib.org/dlib/June97/061lagoze.html [Accessed 2nd Aug 2009].

Luhn, H. P. (1957) A statistical approach to mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, 1, p.309-317.

Patitungkho, K. and Deshpande, N. J. (2005) Information seeking behavior of faculty members of Rajabhat University in Bangkok. *Webology* [online] 2(4). Available from: http://www.webology.ir/2005/v2n4/a20.html [Accessed 30th Dec 2010].

Saracevic, T. (1996) Modeling interaction in information retrieval (IR): a review and proposal. *Proceedings of the ASIS Annual Meeting*, 33, p.3-9.

Saracevic, T. (1997) The stratified model of information retrieval interaction: Extension and applications. *Proceedings of the American Society for Information Science*, 34, p.313-327.

Schneiderman, B. et al. (1997) Clarifying search: A user-interface framework for text searches. *D-Lib Magazine*, January 1997. [Online], Available: http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html [Accessed 2nd Aug 2009].

Spink, A et al. (2002) Information-Seeking and mediated searching. Part 1. Theoretical framework and research design. *Journal of the American Society for Information Science and Technology*, 53(9), p.695-703.

Spink, A. et al. (2001) Searching the Web: the public and their queries. *Journal of the Americam Society for Information Science and Technology*, 52(3), p.226-234.

Spink, A. (1999) Toward a theoretical framework for information retrieval within an information-seeking context. *Proceedings of the 2nd International Conference on Information Seeking in Context*, p.12–15.

Sullivan, D. (2009) Search engine sizes. [Online], Available: http://searchenginewatch.com/reports/sizes.html [Accessed 2nd August 2009]

Sutcliffe A. and Ennis, M. (1998) Towards a cognitive theory of information retrieval. *Interacting with computers*, 10(3), p.321-351.

UMLS (1994) UMLS knowledge source. 5th Experimental Edition. Bethesda, MD: National Library of Medicine.

Vakkari, P. (1999) Task complexity, information types, search strategies and relevance. *Integrating studies on information seeking and retrieval*, p.35-54.

Wheatley, H. B. (2002) What is an Index? A Few Notes on Indexes and Indexers (p.28). London: Longmans, Green & Co.

Wilson, T. D. (2000) Human Information behavior. *Informing Sciences,* 3(2), p.49-55.

Wilson, T. D. (1999) Models in information behavior research. *Journal of Documentation,* 55(3), 249-270.

Yates, R. B. and Neto, B. R. (1999) Modern information retrieval. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Chapter 2 – Literature Review

2.0 Introduction

The Internet has become one of the most important and integral information sources for human lives and work. As the use of Internet is accelerated, the information environment has become more complex. World Wide Web search engines became an abundant tool for efficient information access. The user population of search Web is extremely diverse consisting of computer novices and highly skilled experts; some searchers looking for material just for fun and users requiring accurate and efficient search facilities for professional and academic purposes.

Query formulation is an essential part of successful information retrieval. The challenges in formulating effective queries are emphasized in the Web information search, because the Web is used by a diverse population varying in their level of expertise. To resolve the complexities of Web searching, researchers have endeavored to understand how people search for information on the Web in order to serve the user community. Studies on Web search behavior appeared as early as 1995 and have proliferated since (Hsieh-Yee, 2001).

Hence, investigating the query formulation in Web search for a variety of user groups has become a popular topic for the past literature. A review of literature for this study mainly covered the following areas:

- Studies on information searching strategies (or query formulation) on Web searching
- Web information searching behavior

- Web information retrieval
- The way of view of Web results by users
- Problems faced in information searching on Web

2.1 Information searching strategies (or query formulation) on Web searching

Query formulation is an essential part of successful information retrieval. There are currently very few studies focusing on query formulation in Web, whereas the literature on traditional search environments is plentiful. Furthermore, despite the differences in search environments, in all text based information retrieval the underlying problem is the same for the user: how to communicate the information needed to the computer so that relevant information is retrieved.

Aula, et *al*. (2005) conducted a study on "Modeling Successful performance in Web search". Instead of dividing the searchers into different groups on their expertise, they chose to model search success with Task Completion Speed (TCS). Towards this goal, 22 participants performed three fact-finding tasks and two broader tasks as an observational user study. In the study there were two variables related to the Web experience of the participants. The other variables were the speed of query iteration, the length of the queries, the proportion of precise queries, and the speed of evaluating result documents. Results showed that the variables related to Web experience had effects on TCS. The increase in the years of Web use was related to improvement in TCS in the broader tasks, whereas the less frequent Web use was related to decrease in TCS in the fact-finding tasks. They further found that other variables having significant effects on TCS in either of the task types were the speed of composing queries, the average number of query terms

per query, the promotion of precise queries, and the participant's own evaluation of their search skills.

From the search problem through query formulation to results on the web was done by Barsky and Bar-Ilan (2005) to investigate the effectiveness of the instructions. The experiment took place at the School of Library, Archive and Information Studies at the Hebrew university of Jerusalem. A total of 38 first year MLIS student participated in the experiment. The experiment consisted of two exercises. The student didn't receive comments on the first exercise until the end of the experiment. After the first exercise, a lecture was delivered on the internet search instruction. The difference between the total averages (48.86 before the lecture and 56.51 after the lecture) is significant according to the t-test: p=0.041. They can conclude that, as a result of the lecture, students significantly improved their results in the second exercise.

A study was conducted by Barsky *et al.* (2005) with the purpose of creating Internet search instructions, to test their effectiveness and to tract the search behavior of the first year MLS students. Results revealed that as a result of the lecture students significantly improved the search results. There were a certain number of problematic search behavior patterns and obstacles uncovered, which proved to have a critical impact on the search results; such as difficulties in acquiring new of alternative vocabulary during the search process and perception of the task as the ultimate source of keywords etc.

'Query formulation in Web information search' was done by Aula (2003). This study the author considered as the base for his research. This survey was conducted mainly for students and staff of the University of Tampere and data collected through questionnaires.

The questionnaire consisted of four main parts: the instructions, 20 search task, topic familiarity evaluation part, and ten background questions. In the questionnaire, the participants were asked to formulate the initial queries for the given search tasks. He examined that the average number of search terms was 3.0 per query and there was a positive correlation between Web experience and the average length of the formulated queries. The correlation between the frequency of using computers and the average number of query terms per search was also statistically significant since Web and computer experience are closely related. The results of the study suggested that experience in using computers; Web and Web search engines affect the query formulation process. Further they found that experienced users formulate longer and more specific queries, where as queries of users with less experience of use fewer and more generic terms. The expertise in the use of the tools needed for information search (computers, the Web, search engines) and the expertise in the domain of the search is also known to affect search behavior. People who are familiar with the domain of the search can be expected to use high-quality search terms along with appropriate synonyms; they even claim to be more systematic searchers who plan their searches beforehand (Hölscher and Strube, 2000; Jenkins et al., 2003; Navarro-Prieto et al., 1999). However, to benefit from the high domain expertise seems to require some experience with the search system (Hsieh-Yee, 1993; Vakkari et al., 2003). Generally, the searchers who can rely on both high domain and tool expertise complete search tasks most fluently (Hölscher and Strube, 2000; Jenkins et al., 2003).

2.2 Web information searching behavior

Information searching or Information Retrieval (IR) has become one of the "hottest" research areas in information studies during the last decades. Many Web search engines are based and developed on the basis of the results of these studies and IR principles. Jansen *et al.* (1999) argues, however, that the Web searching is changing IR considerably, and that Internet IR is a different IR, with a number of implications.

A study of "Web search behavior of university students: a case study at University of the Punjab" carried out by Malik and Mahmood (2009) explored the different aspects of web search behavior of university students, in terms of users' background and experience with web, purpose of use, searching skills, query formulation, frequency of use, favorite search engine, etc. of the students of the Faculty of Economics and Management Science. Results revealed that the respondents used basic search frequently with a mean of 3.86 on a 1-5 scale; where 1 meant rarely and 5 meant often, advanced search moderately with a mean of 3.11 and a little use of web directory (mean: 2.5) was made for the searching of materials. A frequent trend of multiple-query search has also been denoted. As the quality of web resources, students were asked whether they found relevant information on the web. The results stated that to some moderate extent, they usually found what they need. The responses demonstrated that students usually identified the relevancy by title and highlighted words. Descriptors and reading URLs are less used.

White, *et al.* (2009) carried out a study on "Characterizing the influence of domain expertise on web search behavior" with the aim of studing the web search behavior of different domain expertise (Medicine, Finance, Law and Computer Science). The study

was a large-scale, longitudinal, log-based analysis of the effect of domain expertise on web search behavior. Results revealed that the sessions conducted by domain experts were generally longer than non-experts' sessions. Domain experts consistently visited more pages in sessions, an in three of the four domains they spent more time and post more queries. Further, they found experts' sessions are more diverse than non expert sessions.

Nicholas *et al* (2009) carried out a study on "student digital information seeking behavior in content" with the purpose of studing the actual information seeking behavior of students in digital scholarly environment. They used log data from two digital journal libraries, Blackwell Synergy and Ohio LINK, and one e-book collection (Oxford Scholarship Online). The study showed a distinctive form of information seeking behavior associated with students and the differences between them and other members of academic staff. For example, students constituted the biggest users in terms of sessions and pages viewed, and they were more likely to undertake longer online sessions. Further, when searching they found that students were however more likely to undertake a simple search.

Zhang *et al* (2005) conducted an exploratory study of engineering and science students to see how domain knowledge affects users' search behavior and search effectiveness. Their study concludes that the level of domain knowledge seems to have an effect on search behavior (that is, as this level increases, the user tends to do more searches and to use more terms in queries), but not on search effectiveness.

Web search engines face an extremely heterogeneous user population from web novices to highly skilled experts. About 85% of the users of the World Wide Web have been estimated to use search engines (Kobayashi and Takeda, 2000). However, novices are shown to have considerable difficulties in finding information from the web (Pollock and Hockley, 1997), while experts are more successful. Few studies have been done to compare the expert and novices behavior of web searching.

A study conducted by Aula *et al* (2003) on "Understanding experts search strategies for designing user friendly search interfaces" was done in order to study the search strategies of the experienced web searchers. Seven computer scientists were observed during and interviewed after performing their own work-related web search tasks. The results revealed the experts have effective means for enhancing searching, such as: using simple search terms & operators, frequent query editing, using multiple windows, versatile results saving, and using the "find" functionality.

Web searchers differ from each other in many ways that can greatly influence their ability to carry out successful searches. One way in which they can differ is in their knowledge of a subject or topic area. Previous research suggests that domain experts employ different search strategies and are more successful in what they are looking for, than non-experts. Research on domain expertise has examined differences between experts and non-experts in three main classes of search behavior: Query attributes (choice of search terms, query length and syntax), search strategies and tactics (Resource selection, sequence of actions, mix of querying and browsing), and search out comes (accuracy, time).

Jansen *et al* (2000) carried out a study on "Real life, real users, and real needs: a study and analysis of user queries on the Web" in order to study the Web searching behavior. They analyzed transaction logs containing 51473 queries posted by 18113 users of *Exite,* a major Internet search service. They found that most users did not have many queries per search. The mean number of queries per user was 2.8. Further, they found that Web queries are short. On average, a query contained 2.21 terms. They further revealed that Boolean operators are seldom used. One in 18 users used any Boolean operator. As to the queries, about one in 12 queries contained a Boolean operator, and in those "AND" was used by for the most. They found that the ability to create phrases (terms enclosed by quotation marks) was seldom, but correctly used. On average, users viewed 2.35 pages of results (where one page equals ten hits). Over 50% of the users did not access results beyond the first page.

White, Domas and Livonen, (2001) carried out a study on "Questions as a factor in Web searching strategy" in order to study the initial Web searching strategies of American and Finnish students. They found that overall; the participants relied most heavily on search engines (43%) then on direct address (30%) and to slightly lesser extent on directories (27%).

Another study carried out by Spink *et al* (1999) on "Searching the Web: the public and their queries" and found that most people use few search terms, few modified queries, view few Web pages and rarely used advanced search features. Results revealed that less than 5% of all queries used any Boolean operators. Of these AND was most used. A smaller percentage of queries used OR and a minuscule percentage used AND NOT. Further, it revealed that the mean number of terms in unique queries was 2.4. They found

that a great majority of Web queries posed by the public is short, not much modified, and very simple in structure.

2.3 Number of pages viewed per user

Not much research has been done to examine the results viewing patterns of Web users**.** Web users view few result pages; they tend not to browse beyond the first or second page of results. Malik, Amara and Mahmood, Khalid (2009) in their study "Web search behavior of university students: a case study at University of the Punjab" revealed that most of the students (56%) generally browsed through the first ten result page of the web. Only 26.5% students browsed 20 hit results.

Jansen and Spink (2003) conducted a research on the "Analysis of Web documents retrieved and viewed". The results revealed that 53% of the users entered one query and about 54% of the users viewed only one page of results. They further found that the relationship between the number of queries submitted and the number of results pages viewed parallels each other with about equal percentages of queries submitted and results pages viewed.

A study carried out by Spink, *et al* (1999) on "Searching the Web: the public and their queries" showed the median was 8 pages viewed per user. However, 28.6% of users examined only one page of results, i.e. since a page contains ten ranked Web sites, about one in every four users looked at ten or less sites. Another 19% looked at two pages only. That is, close to half of the users looked at two or less pages.

2.4 Summary of the Literature Review

Studies focusing on information search strategies (or specifically query formulation) have studied professionals searching from pre-web information retrieval systems (Mansourian, 2004; White, 2001; Iivonen and Sonnenwald, 1998). These studies present the background for studying query formulation in web searching. People have trainings in query formulation on pre-web information retrieval systems. In the web environment the trainings are very lack. In this regarding the search engines play a vital role in Web. They have been designed to serve the expertise and the novices as well, via query formulation. There are currently very few studies focusing on query formulation in Web, whereas the literature on traditional search environments is plentiful. Furthermore, despite the differences in the search environments, in all text based information retrieval the underlying problem is the same for the user: how to communicate the information need to the computer so that relevant information is retrieved.

Studies of web information searching behaviour have usually focused on very large search engine logs files (e.g., Anick, 2003; Jansen and Pooch, 2001; Jansen et al., 2000; Silverstein et al., 1999; Spink et al., 2000). Most of these studies are based on quantitative data. In general, the log studies have shown that web searchers use short queries (typically from 1 to 3 terms), seldom use advanced operators, do not regularly iterate their queries, and only go through a couple of result pages per query (Anick, 2003). There were some studies comparing the differences in the search behavior of novices and experts (Hölscher and Strube, 2000; Sutcliffe et al., 2000). According to those studies in query formulation experts typically use longer queries and the advanced operators than novices. The study done by Vakkari (2000) shows that more detailed and longer queries are

formulated by people who are more familiar with the topic. As shown in the study of White and Iivonen (2001), the search task also affects the search strategies. The open vs. closed nature of the task has shown to affect novices and experts differently. The study by Silverstain et al (1999) showed that almost 64 % of the search sessions in their data set (log data from AltaVista search engine) consisted of only one query. However, the reason for this could not be inferred from the data: It is possible that the users found the information they were looking for immediately or they possibly gave up as soon as they noticed that the search was not successful. Belkin (2000) has aptly illustrated the challenge the users face in text-based information retrieval: "How to guess what words to use for the query that will adequately represent the person's problem and be the same as those used by the system in its representation." For some users, this task is presumably easier than for others and our goal is to study the user characteristics affecting the "guesses" they make.

Some authors have defined the terms 'expert' and 'novice' and characterized the differences between these groups. Most frequently, Web experience (either the years of experience or the frequency of usage) has been used in dividing the users into novices and experts. However, the requirements for the experience levels have differed considerably between the studies. Thus, the "experts" in one study could easily be "novices" in another. This ambiguity in defining experience has certain consequences. First of all, the results from different studies cannot be directly compared and hence, the studies do not add to the understanding of the characteristics of the different groups. It has the consequence that differences may rather be due to very different groups of users being compared.

The queries of experienced users tend to be longer than those of the less experienced users (Aula, 2003; Hölscher and Strube, 2000), the use of Boolean operators and modifiers appears to be higher among "professional" users (Aula and Siirtola, 2003; Hölscher and Strube, 2000), and the experienced users are found to commonly iterate their queries when the information is not immediately found (Aula and Käki, 2003). Novices (or typical search engine users) on the other hand, frequently make errors in formulating sophisticated queries (Aula, 2003; Jansen et al., 2000). Navarro-Prieto *et al* (1999) found that in fact-finding tasks, the style of the queries differed between more and less experienced users: experienced users frequently used a bottom-up strategy (the query terms are taken directly from the instructions), whereas the novices used very general queries (similar findings reported in Aula, 2003). Fields et al. (2004) reported that experts (librarians) showed excellent skills in reformulating their queries so that each reformulation was built on the previous query in retrieving results.

The review of literature presents a picture that different areas of Web searching. No study had been conducted on the factors affecting query formulation in web searching among university academic community so far. Thus, this study is the first of its nature conducted in Sri Lanka.

Reference

Anick, P. (2003) Using terminological feedback for web search refinement: a log-based study. *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, p.88-95.

Aula, *et al* (2005) Modeling successful performance in Web [Online], Available from: http://www.cs.uta.fi/reports/dsarja/D-2005-4.pdf [Accessed 25th March 2010].

Aula, A. (2003) Query formulation in web information search. *Proceedings of the IADIS international conference*, p. 403–410.

Aula, Anne and Kaki Mika (2003) Understanding expert search strategies for designing user-friendly search interfaces. [Online], Available from: http://www.iadis.net/dl/final_uploads/200302C098.pdf [25th March 2010].

Barsky and Bar-Ilan (2005) From the search problem through query formulation to results on the web. *Online information review*, 29(1), p.75-89.

Belkin, N. J. (2000) Helping people find what they don't know. *Communications of the ACM*, 43(8), p.58-61.

Fields, B., Keith, S., & Blandford, A. (2004) Designing for expert information finding strategies. *Proceedings of HCI*, p. 89-102.

Hölscher, C., and Strube, G. (2000) Web search behavior of internet experts and newbies. *Proceedings of the 9th International WWW conference*, p. 337-346.

Hsieh-Yee, Ingrid. (2001) Research on Web search behavior. *Library and Information Science Research*, 23, p.167-185.

Hsieh-Yee, Ingrid. (1999) Effects of search experience and subject knowledge on online search behavior: Measuring the search tactics of novice and experienced searchers, *Journal of the American Society for Information Science*, 44(3), p.161-174.

Iivonen, M. and Sonnenwald, D. H. (1998) From Translation to navigation of different discourses: A model of search term selection during the pre-online stage of the search process. *Journal of the American Society for Information Science*, 49(4), p.312-326.

Jansen, B.J., et al (1999) The Use of Relevance Feedback on the Web: Implications for Web IR System Design. *Proceedings of WebNet World Conference on the WWW and Internet*, p.550-555.

Jansen, B.J., et al (2000) Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing and Management*, 36(2), p. 207-227.

Jansen, B.J. and Pooch, U. (2001) Web user studies: a review and framework for future work. *Journal of the American Society for Information Science and Technology*, 52(3), p. 235-246.

Jansen, B. J. and Spink, A. (2003) An analysis of Web documents retrieved and viewed. *The 4th International Conference on Internet Computing*, p.65-69.

Jenkins, C., Corritore, C.L., & Wiedenbeck, S. (2003) Patterns of information seeking on the Web: A qualitative study of domain expertise and Web expertise. *IT & Society*, 1(3), p.64-89.

Kobayashi, M. and Takeda, K. (2000) Information retrieval on the Web. *ACM Computing Surveys*, 32(2), p. 144-173.

Malik, Amara and Mahmood, Khalid (2009) Web search behavior of university students: a case study at University of the Punjab, *Webology*, 6(2), p.1-14.

Mansourian, Yazdan (2004) Similarities and differences between web search procedure and searching in the pre-web information retrieval systems. [Online], Available from: http://www.webology.ir/2004/v1n1/a3.html [Accessed 3rd March 2009].

Navarro-Prieto, R., Scaife, M., and Rogers, Y. (1999), Cognitive strategies in Web searching. *Proceeding of the 5th Conference on Human Factors & the Web*. [Online], Available: http://zing.ncsl.nist.gov/hfWeb/proceedings/navarro-prieto/. [Accessed 7th Aug 2009].

Nicholas, D., Huntington, P., Jamali, H. R., Rowlands, I., and Fieldhouse, M., (2009) Student digital information-seeking behavior in context, *Journal of Documentation*, 65(1), p.106-132.

Pollock, A. and Hockley, A. (1997) What's wrong with Internet searching. *D-Lib Magazine*, Available: http://www.dlib.org/dlib/march97/bt/03pollock.html. [Accessed 7th Aug 2009].

Silverstein, C. *et al.* (1999) Analysis of a very large web search engine query log. *SIGIR Forum*, 33(1), p. 6-12.

Spink, A., Bateman, J. and Jansen, B. J. (1999) Searching the Web: a survey of EXCITE users. *Internet Research*, 9(2), p.117-128.

Spink, A., Jansen, B. J., and Ozmultu, H. C., (2000) Use of query reformulation and relevance feedback by Excite users. *Internet Research*, 10(4), p.317- 328.

Sutcliffe, A.G., Ennis, M., and Watkinson, S.J. (2000) Empirical studies of end-user information searching. *Journal of the American Society for Information Science*, 51(13), p. 121-123.

Vakkari, P. (2000) Changes in relevance criteria and problem stages in task performance. *Journal of Documentation*, 56(5), p.540-562.

Vakkari, P., Pennanen, M. and Serola, S. (2003) Changes of search terms and tactics while writing a research proposal: a longitudinal case study, *Information Processing and Management*, 39(3), p. 445-463.

White, Marilyn Domas, and Iivonen, Mirja (2001) Questions as a factor in Web search strategy. *Information Processing and Management*, 37, p.721-740.
White, R.W., Dumais, S.T., and Teevan, J., (2009) Charactering the influence of domain expertise on web search behavior. *Proceeding of the second ACM international conference on Web search and data mining*, p.132-141.

Zhang, X., Anghelescu, H., and Yuan, X., (2005) Domain knowledge, search behaviour, and search effectiveness of engineering and science students: an exploratory study, *Information Research*, 10(2) p.217.

Chapter 3 – Research Design

The survey design was used for this study as it allows the researcher to gather information about a target population without undertaking complete enumeration. The study was conducted as a case study since an in-depth analysis on factors affecting query formulation in Web information searching was to be carried out. The case studies allowed the researcher to obtain comprehensive information about the research objectives and also to obtain a comprehensive examination of problems unique to the organization and relationships between individual groups (Yin, 1994).

3.1 Population

The target population for the research was the students who accessed Internet for academic and research purposes. The University of Moratuwa provides Internet access at different levels for different user categories. Academic staff members could access the Internet personally in their respective departments. Students also could access the Internet at their respective departments, as well as in the library. As both academic staff members and the students of University of Moratuwa have accessibility to the Internet,

- Academic staff members,
- Postgraduate students,
- Undergraduate students and
- Students of National Diploma in Technology (NDT)

were the category wise elements of population selected for this research.

3.1.1 Exclusive criteria

First year undergraduates, NDT students and librarians were not included in the target population of the study. As new comers to the university, the first year undergraduates and NDT students are not much aware of facilities (such as Internet) provided at the university. Normally the students were assigned to a particular department after the end of the first year examination. Therefore, the first year undergraduates did not belong to a specific department and they had very limited facilities to browse the Internet. A less number of librarians (07) were working in the university library and their subject disciplines totally differed from other academic staff members' of the University of Moratuwa. Therefore, they couldn't be included in the analysis. So librarians were not conceded for the target population of the study. Since there weren't any other post graduate level graduate courses for NDT students, nobody was further added to the target population from NDT face.

3.1.1 Inclusive criteria

All undergraduates and NDT students, (except first year students), all postgraduate students, and all academic staff members (except librarians) were the elements of the target population.

The students' registration list at the Examination Branch and academic staff registration list at the Academic and Establishment Division were the source used in defining students and academic population. A detail stratification of the population was presented in Table 3.1.1, Table 3.1.2 and Table 3.1.3.

Table 3.1.1: Distribution of undergraduate and Diploma students by faculties / University of Moratuwa

Faculty / Institute	Year 2	Year 3	Year 4	Year 5	Total
Architecture	298	293	261	306	1158
Engineering	783	739	771		2293
Information Technology	104	98	101		303
Institute of Technology	322	297	-	-	619
Total	1507	1427	1133	306	4373

Source: Examination & Academic Division University of Moratuwa
Institute of Technology, University of Moratuwa, (September 2010)

Table 3.1.2: Distribution of Postgraduate students by faculties/ University of Moratuwa

Faculty	No of Pg students
Architecture	164
Engineering	734
Information Technology	60
Total	**958**

Source: Examination & Academic Division University of Moratuwa (September 2010)

Table 3.1.3: Distribution of academic staff by faculties / University of Moratuwa

Faculty	Senior Prof.	Prof.	Ass. Prof.	Senior Lecture Grade I	Senior Lecture Grade II	Lecture (Probationary)	Total
Architecture		5	2	12	19	18	56
Engineering	7	25	1	37	57	62	189
Information Technology		1		2	6	9	18
NDT				2	11	27	40
Total	**7**	**31**	**3**	**53**	**93**	**116**	**303**

Source: Establishment Division University of Moratuwa
Institute of Technology, University of Moratuwa, (September 2010)

3.2 Sample Selection Criteria

The sampling technique used for the study is “stratified random sampling”, since respondents were randomly selected from each stratum for representation in the sample. This kind of sampling allows comparing characteristics between different strata.

The target academic staff population was not sufficient to be considered as faculty and academic position, classified in Table 3.1.3. Therefore, they were taken as one group comprising i.e. academic staff. Postgraduate students were divided according to their faculties. Undergraduate & NDT students were divided according to their faculties /institute and year of study.

The sample size was determined according to Yamane (1967) simplified formula under 5% precision level and 95% confidence level.

Yamane simplified formula is n = $\frac{N}{1+N(e^2)}$

Where, n= Sample size, N= Size of the target population (5634) and e= Precision coefficient (0.05)

According to the above Yamane formula, the sample size would be 373 (373.4836). Appropriate proportions of the sample can be summarized as follows:

Table 3.2.1: Proportions of the sample according to the user categories

User Category	No. of User in the Population	No. of User in the Sample
Undergraduate	3754	249
NDT	619	41
Postgraduate	958	63
Academic Staff	303	20
Total	**5634**	**373**

Table 3.2.2: Classified proportions of the sample/ Undergraduate & Postgraduate students

Faculty / Institute	Undergraduate		Postgraduate	
	Pop	Sample	Pop	Sample
Architecture	1158	**77**	164	**11**
Engineering	2293	**152**	734	**48**
Information Technology	303	**20**	60	**4**
Institute of Technology	619	**41**	-	-
Total	4373	**290**	958	**63**

Table 3.2.3: Classified proportions of the sample/ Academic staff members

	Population	Sample
Academic Staff	303	20

In this classified sampling frame, number of undergraduate students in the faculty of Information Technology (20), postgraduate students in the faculty of Architecture (11) and Information Technology (4) and academic staff members were not sufficient to conduct a statistical analysis. According to the Central Limit Theorem there exists one rule of thumb, in most practical situations; a sample of size 30 is satisfactory (Daniel, 2000, p.130). Therefore, a sample of 30 from each stratum was taken from each category without considering the proportion. This method is considered more suitable to maintain the accuracy of the conclusions drawn on statistical analysis. With this adjustment, the total sample size was 438. The selected sample for the study is displayed in table 3.2.4 and 3.2.5.

Table 3.2.4: Adjusted proportions of Undergraduate and Postgraduate students in the sample

Faculty / Institute	Undergraduate		Postgraduate	
	Pop	Sample	Pop	Sample
Architecture	1158	77	164	30
Engineering	2293	152	734	48
Information	303	30	60	30
Institute of	619	41	-	-
Total	4373	300	958	108

Table 3.2.5: Adjusted proportions of Academic staff members in the sample

	Population	Sample
Academic staff	303	30

3.3 Research Instruments

3.3.1 Survey method

The questionnaire **(Appendix I)** consisting of four parts was selected as a research instrument to collect data.

Part I: Background Information

Factual seven questions were included in this in order to gather some personal and professional data of community of the university.

Part II: Level of computer usage

Four questions were included in this section to gather data to find out the level of computer experience with respondents.

Part III: Status of Web searching

Fourteen questions were included in this part. Participants were asked questions about:

- Frequency of using Web & search engines
- Search engine & Web experience
- Purpose of Web used & type of Web information sources used
- Own rating of search skills
- Usage of advance searching techniques
- Courses or Web searching guides on information retrieval

Part IV: Search task – Query Formulation

Five search tasks are included in this part and three of the tasks are fact-finding and the remaining two tasks are exploratory (open-ended).

Five search tasks are given per each user and he/she has to search and write down:

i. Topic familiarity
ii. Search statement he/she tried
iii. Time spent on each search task (maximum three attempts allowed tried each task)
iv. If they used advance features of the search engine and how they judged the relevancy of information.

Three questionnaires were set to gather information from undergraduate, postgraduate students and academic staff members. In part I, question numbers 2 and 3 differed for each questionnaire and other questions were the same.

Questionnaires were hand delivered to the selected users since they were easily approachable within the university premises. Email and phone calls were used to follow up with the respondents.

3.3.2 Observation method

Ten elements of each group of the population (undergraduate students, postgraduate students, NDT students and academic staff members) were selected as the observation sample to be convenient for the study. The sample size was 40.

The Observation method other than the questionnaire was used in the study for the result evaluation part. The Observation method was adopted to back check whether the answers provided for five search tasks by respondents, were corrected. Descriptive analysis of this observational data was used to confirm the answers written in the questionnaire on search tasks. Observations were made to check whether the searchers have tended to attempt 2nd or 3rd turn of query reformulation or whether they have become satisfied with results of the 1st query criteria formulated by them.

3.4 Evaluation of results

The user's task in the result evaluation phase was to decide whether the individual results are worth a more detailed inspection and whether the query needs to be refined. In essence, the user needs to evaluate the relevance of the documents retrieved (Aula, 2005). Relevance is a more complex concept to define. Mizzaro (1997), in his thorough review

of the history of relevance, presents a framework for various kinds of relevance. The framework is too complex. Therefore, in most studies the complexity of relevance has been simplified by making binary relevance judgments: each document is simply judged relevant or not relevant. The relevance is commonly related to the topic of the search: if the document deals with the topic (at all), it was considered relevant (Aula, 2005). This method was used for the present study. Researcher identified the key words and synonyms for each search task prior to distributing the questionnaire and observed whether results included all of them; then it was taken as being relevant.

Usually, web search engine users make their initial evaluation of the relevance of the results by relying on a textual representation of the underlying document or document surrogate (Veerasamy & Heikes, 1997). In this study, the surrogates are also referred to as search results. Search engine users tend to check only the first result page (10 results) for each query (Jansen & Pooch, 2001).

To measure precision in Web searches (relevance results as a proportion of the number of results retrieved), the calculation was taken into account as follows for the current study;

Precision = (no. of relevant documents in the first ten results)/10*100

3.5 Treatment of Data

SPSS (Statistical Package for Social Scientists) 13.0 for Windows was used to occupy a variety of statistical techniques. The analyzed data is represented in forms of tables, bar-charts, pie-charts, etc. The following statistical tests were performed to analyze the result of the survey.

Chi-square test

Chi-square test was performed to test the association between variables over the hypothesis.

H_0: two variable are independent verses H_1: two variable are dependent

Kruskal-Wallis test

Kruskal-Wallis One-way Analysis of Variance test (Non-parametric test) was performed to check the equality of medians for two or more populations. The Kruskal-Wallis hypotheses were:

H_0: the population medians are all equal verses H_1: the medians are not equal

Pearson's correlation coefficient

Pearson's correlation coefficient is a measure of linear association of two variables.

Friedman test

Friedman test is a nonparametric analysis of a randomized block experiment, and thus provides an alternative to the Two-way analysis of variance. The testing hypotheses are:

H_0: all treatment effects are zero versus H_1: not all treatment effects are zero

3.6 Limitations and Delimitations of the Study

As limitation and delimitations are an integral part of the study, they affect the design, as well as the outcome of the study.

The factors beyond the control of the research would affect the research and outcomes of the research. Those limitations and delimitation could be listed as follows:

- Undergraduate, NDT, and postgraduate students registered at the University of Moratuwa (as at the year 2010) and academic staff members who participated in the study.
- The exclusion of the first year undergraduate and NDT students who did not participate in the study because they did not have enough facility to browse internet in the University when compared with other students.
- Librarians did not participate in the study. Their number (07) was very few. They cannot be considered with other academic staff members, since their disciplines differed from other academic staff.
- When elements of the classified sampling frame was less than 30, the sample size was taken as 30 elements for each stratum.

3.7 Operational Definitions

Information Retrieval

The technique and process of searching, recovering, and interpreting information from large amounts of stored data.

Search Strategy

A systematic plan was used to conduct a search. In most cases, the first step is to formulate a clear and concise topic statement. The next step is to identify the main concepts in the topic. Next, the most appropriate finding tools for the subject must be identified and located.

References

Aula, A. (2005) Studying user strategies and characteristics for developing web search interfaces. Academic dissertation. University of Tampere.

Daniel, Wayne W. (2000) *Bio statistics: a foundation for analysis in the health science*. John Wiley & Sons Inc. p. 130.

De Silva, K.M. and Peiris, G.H. (1995) The University system of Sri Lanka. International Center for Ethnic Studies , Kandy.

Glenn, D. Israel. Determining sample size. University of Florida IFAS Extension. Available from: http://edis.ifas.ufl.edu. [Accessed 2nd Nov 2010]

Jansen, B.J. and Pooch, U. (2001) A review of web searching studies and a framework for future research. *Journal of the American society for Information science and technology* 52(3), p.235-246.

Karunaratne, S. and Dias W.P.S. (1997) University of Moratuwa: History 1972-1997. University of Moratuwa, Moratuwa.

Mizzaro, S. (1997) Relevance: the whole history. *Journal of the American Society for information Science* 48(9), p.810-832.

University of Moratuwa (2010) University handbook 2010. University of Moratuwa, Moratuwa

Yin, Robert K. (1994) *Case study research design and methods*. 2nd edition. New Delhi, SAGE publications Ltd.

Chapter 4 – Analysis, Interpretation and Discussion

Analysis, its interpretations and the discussion are presented in this chapter in order to gradually achieve the objectives of the study. The analysis is based on the gathered data from the questionnaires.

4.1 Description of the sample

The sample data is described in this section.

4.1.1 The response rate

The response rate is important in a study since the validity of any statistical analysis depends on the extent of the response rates of various categories of the target population. Out of 438 questionnaires, 432 were returned at a rate of 98.63%. Table 4.1.1 shows the details of the response rate.

Table 4.1.1: Response rate of each user group

User Group	Sample Size	Number of Respondents	Response Rate
Academic Staff Members	30	30	100.00%
Postgraduate Students	108	107	99.07%
Undergraduate Students	259	255	98.46%
NDT Students	41	40	97.56%
Total	**438**	**432**	**98.63%**

According to table 4.1.1, the highest response rate is observed with the academic staff members (100%) and the lowest response rate is observed for NDT students of the University of Moratuwa. The response rates of postgraduate and undergraduate students are 99.07% and 98.46% respectively. However, it can be concluded that overall response rate (98.63%) is satisfactory to draw conclusions for the target population.

4.1.2 Distribution of respondents

The respondent's profile is described in terms of their gender, age, ethnicity and academic position for each user group.

Distribution of the respondents among the faculty/ institute according to gender is summarized in table 4.1.2.1. It shows that the percentages of the male respondents were higher for each faculty. Table 4.1.2.1 further depicts that the gender distribution is not much varied over the respondents in the Faculty of Architecture. It is categorized as 52.2% males and 47.8% females.

Table 4.1.2.1 Distribution of respondents according to gender and faculty/ institute

Faculty/Institute	Male	Female	Total
Architecture	**59 (52.2%)**	54 (47.8%)	113
Engineering	**160 (75.8%)**	51 (24.2%)	211
IT	**39 (60.9%)**	25 (39.1%)	64
NDT	**37 (84.1%)**	7 (15.9%)	44
Total	**295 (68.3%)**	137 (31.7%)	432

Figure 4.1.2.1 summarizes the distribution of respondents by gender for each user group. It also shows that the highest percentages are comprised of males. The percentages are; 90% NDT students; 67.45% undergraduate students; 63.55% postgraduate students and 63.33% academic staff members. The difference between male and female is higher for the NDT students than others. It may be because the male students normally tend to follow the NDT course. Accordingly, the majority of the students and staff members are males in the sample.

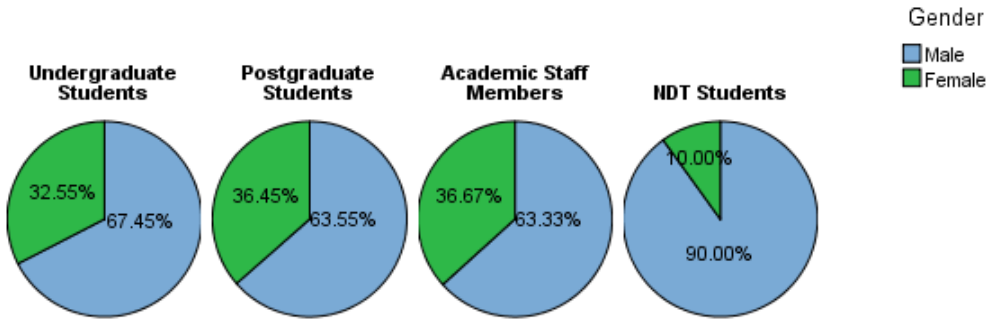


Figure 4.1.2.1 Respondents by gender

Table 4.1.2.2 further describes the distribution of respondents according to their age groups. It shows that most of the undergraduate (100%) and NDT (97.50%) students are in 20-30 years age group. 44.86% and 41.12% postgraduate students are in 20-30 and 31-40 years age group respectively, and most of the staff members (53.33%) are in 41-50 years age group.

Table 4.1.2.2 Respondents by age

Category	Age (Years)					**Total**
	<20	21-30	31-40	41-50	51<	
Undergraduate Students	0 .0%	255 100.0%	0 .0%	0 .0%	0 .0%	255 100.0%
Postgraduate Students	0 .0%	48 44.9%	44 41.1%	13 12.1%	2 1.9%	107 100.0%
Academic Staff Members	0 .0%	4 13.3%	7 23.3%	16 53.3%	3 10.0%	30 100.0%
ITUM Students	1 2.5%	39 97.5%	0 .0%	0 .0%	0 .0%	40 100.0%
Total	1 .2%	346 80.1%	51 11.8%	29 6.7%	5 1.2%	432 100.0%

Figure 4.1.2.3 presents the distribution of ethnicity of respondents. It illustrates that of the students considered in the sample study; 89.81% are Sinhala, 6.48% are Tamil and 3.70%

are Muslim. In the population, the majority of the students are Sinhala. Therefore, a higher percentage of Sinhala students are represented in the sample.

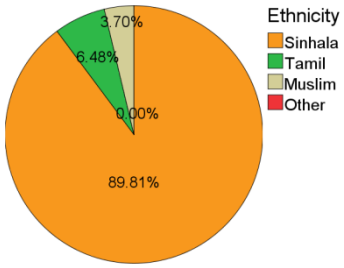


Figure 4.1.2.3 Respondents by ethnicity

4.1.2.1 Undergraduate and NDT students

Out of 295 undergraduate and NDT students who responded to the survey; 77 (26.1%) were from the Faculty of Architecture, 148 (50.2%) were from the Faculty of Engineering, 30 (10.2%) were from Faculty of Information Technology and 40 (13.6%) were NDT students. A total of 255 undergraduate students responded to the survey; and 80 (31.4%) were second year students, 79 (31.0%) were third year students, 76 (29.8%) were fourth year students and 20 (7.8%) were fifth year students. Only 40 NDT students responded to the survey and 50% of them were second year students and the other 50% were third year students. Their field of specialization with respect to the year of study is presented in Figure 4.1.2.1.1.

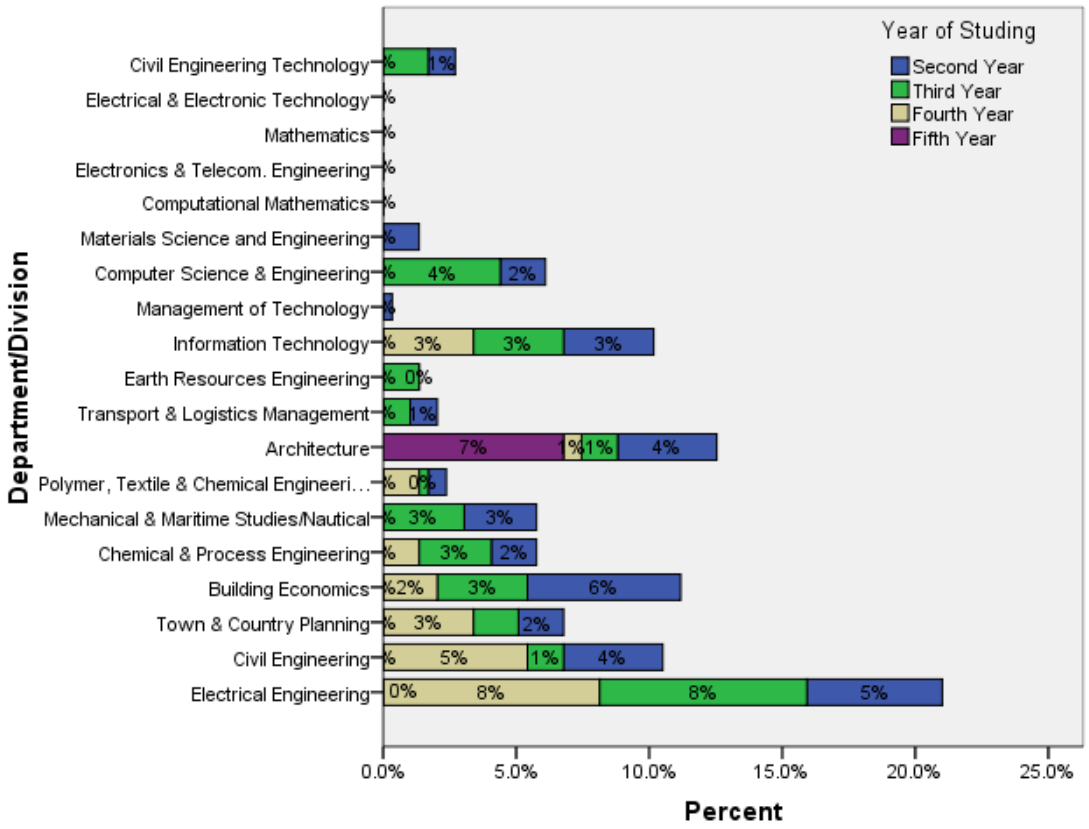


Figure 4.1.2.1.1 Undergraduate and NDT students by department/ division

4.1.2.2 Postgraduate Students

Table 4.1.2.2.1 describes that 43.9% responded postgraduate students were from the Faculty of Engineering. Out of 107 postgraduate students who responded; 78 (72.9%) were reading for M.Sc. degree, 10 (9.3%) reading for M.Eng. degree and 19 (17.8%) reading for MBA degree.

Table 4.1.2.2.1 Postgraduate students by faculty and course

Faculty	Course			Total
	M.Sc.	M.Eng.	MBA	
Architecture	30 (100.0%) (38.5%)	0 (0.0%) (0.0%)	0 (0.0%) (0.0%)	**30** (100.0%) (28.0%)
Engineering	23 (48.9%) (29.5%)	10 (21.3%) (100.0%)	14 (29.8%) (73.7%)	**47** (100.0%) (43.9%)
Information Technology	25 (83.3%) (32.1%)	0 (0.0%) (0.0%)	5 (16.7%) (26.3%)	**30** (100.0%) (28.0%)
Total	**78** (72.9%) (100.0%)	**10** (9.3%) (100.0%)	**19** (17.8%) (100.0%)	**107** (100.0%) (100.0%)

Figure 4.1.2.2.1 shows that the largest percentage of postgraduate students' highest academic qualification was the bachelors' degree (91.43%) and 59.81% were married.

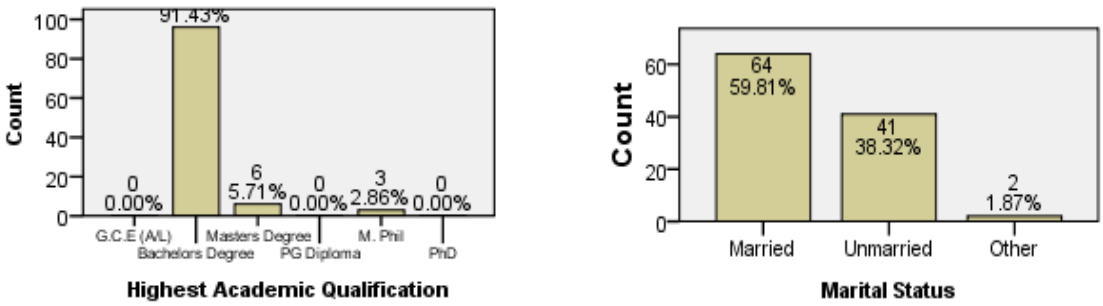


Figure 4.1.2.2.1 Postgraduate students by highest academic qualification and marital status

4.1.2.3 Academic Staff Members

Out of 30 academic staff members respondents; 6 (20%) were from the Faculty of Architecture, 16 (53.3%) were from the Faculty of Engineering, 4 (13.3%) were from the Faculty of Information Technology and 4 (13.3%) were from NDT. The survey included 2 (6.7%) professors, 23 (76.7%) senior lecturers and 5 (16.7%) probationary lecturers.

Table 4.1.2.3.1 Academic staff members by faculty and academic position

Faculty/Institute	Academic Position			Total
	Professor	Senior Lecturer	Lecturer (Prob.)	
Architecture	1 (16.7%) (50.0%)	5 (83.3%) (21.7%)	0 (0.0%) (0.0%)	**6** (100.0%) (20.0%)
Engineering	1 (6.2%) (50.0%)	14 (87.5%) (60.9%)	1 (6.2%) (20.0%)	**16** (100.0%) (53.3%)
Information Technology	0 (0.0%) (0.0%)	2 (50.0%) (8.7%)	2 (50.0%) (40.0%)	**4** (100.0%) (13.3%)
NDT	0 (0.0%) (0.0%)	2 (50.0%) (8.7%)	2 (50.0%) (40.0%)	**4** (100.0%) (13.3%)
Total	**2** (6.7%) (100.0%)	**23** (76.7%) (100.0%)	**5** (16.7%) (100.0%)	**30** (100.0%) (100.0%)

Figure 4.1.2.3.1 summarizes the respondent academic staff members' distribution by highest academic qualification and academic position. It shows that the largest percentage (53.33%) of respondents have a PhD degree and the majority of the respondents were senior lecturers, because the largest group is represented by senior lecturers in the population and most of them have a PhD degree. And also 93.33% respondents of academic staff members were married.

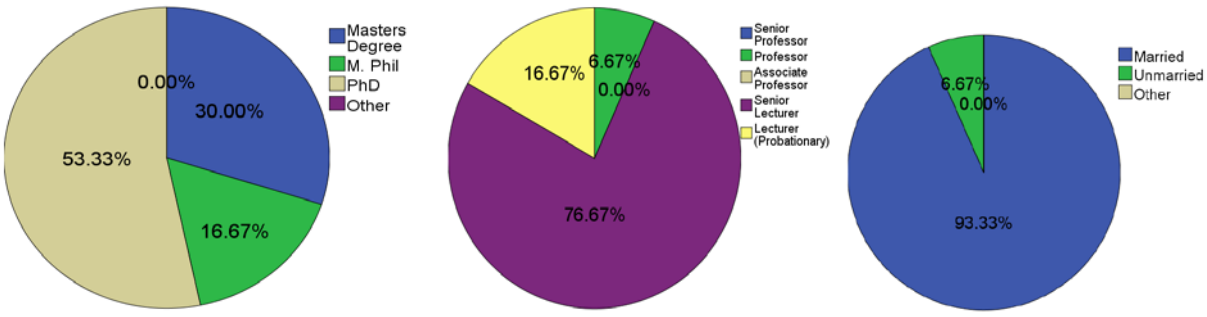


Figure 4.1.2.3.1: Academic staff members by highest academic qualification, academic position and marital status

4.1.3 Level of computer experience of respondents

The following information was gathered under this section:

- Availability of Computer/Laptop/Notebook/PDA and frequency of using it
- Computer courses followed
- Length of computer use
- Availability of internet connection and type of connection

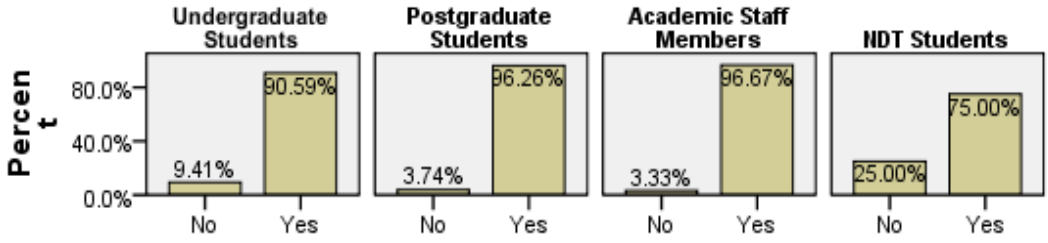


Figure 4.1.3.1 Computer/Laptop/Notebook/PDA availability at home

The above figure 4.1.3.1 presents the computer/laptop/notebook/PDA availability of responders at home. It clearly shows that the majority of the respondents have access to computer/laptop/notebook/PDA at their homes. The relevant percentages are 90.59% for undergraduate students, 75% NDT students, 96.26% postgraduate students and 69.67% academic staff members. 25% of NDT students do not have a computer/laptop/notebook/PDA at home. It is comparatively higher component than the other categories of students with non availability of a computer/laptop/notebook/PDA at home. It may be since their computer applications are not much than other undergraduates.

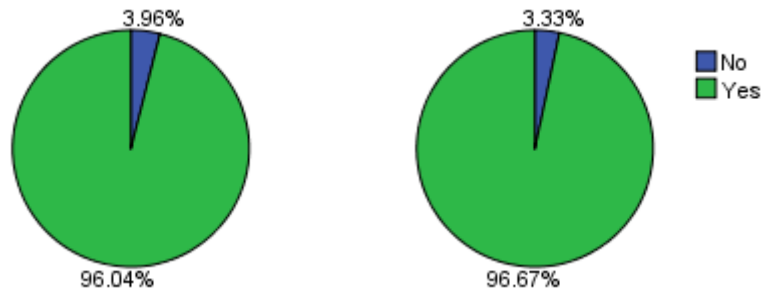
Respondents were also asked the frequency of using a computer/laptop/notebook/PDA at home. Table 4.1.3.1 shows that more than 80% undergraduate students, postgraduate

students and academic staff members are using computers at home daily. But NDT students' daily usage at home is very low comparative to other categories.

Table 4.1.3.1: Frequency of using computer/laptop/notebook/PDA at home

Category	Usage				Total
	Daily	Weekly	Monthly	Less than once a month	
Undergraduate Students	193 83.5%	32 13.9%	1 .4%	5 2.2%	231 100.0%
Postgraduate Students	90 87.4%	12 11.7%	1 1.0%	0 .0%	103 100.0%
Academic Staff Members	24 82.8%	4 13.8%	0 .0%	1 3.4%	29 100.0%
ITUM Students	8 26.7%	20 66.7%	2 6.7%	0 .0%	30 100.0%
Total	315 80.2%	68 17.3%	4 1.0%	6 1.5%	393 100.0%

The distribution of respondents by availability of computer/laptop/notebook/PDA at office is shown in figure 5.1.3.3. It clearly shows that the majority of the postgraduate students (96.04%) and academic staff members (96.67%) have a computer/laptop/notebook/PDA at office.



Postgraduate Students Academic Staff Members

Figure 4.1.3.3 Computer/Laptop/Notebook/PDA availability at office

Figure 4.1.3.4 shows that 98.97% postgraduate students and almost all (100%) academic staff members are using computers daily at their office.

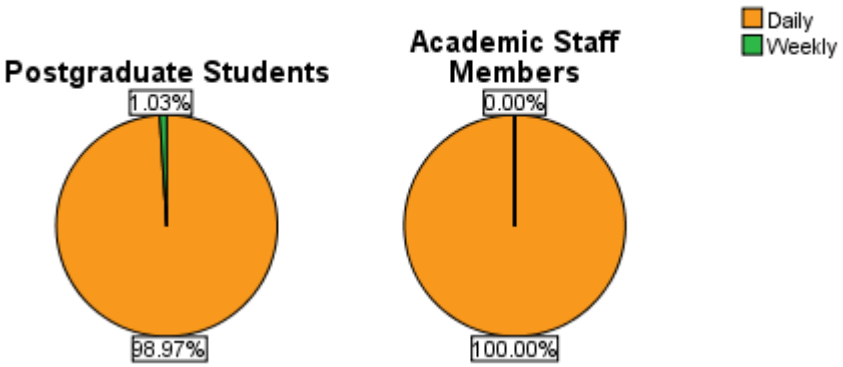


Figure 4.1.3.4 Frequency of using computer/laptop/notebook/PDA at office

Figure 4.1.3.5 summarizes the distribution of respondents who have followed computer course/s by user category. It shows that 76.64% postgraduate students, 68.63% undergraduate students, 65.00% NDT students and 50% academic staff members have followed some computer course. Comparatively, the percentage of postgraduate students who have followed some computer course is higher and the percentage of academic staff members is lower. Post graduate students normally tend to follow computer courses since they have to use computers for most of their academic activities.

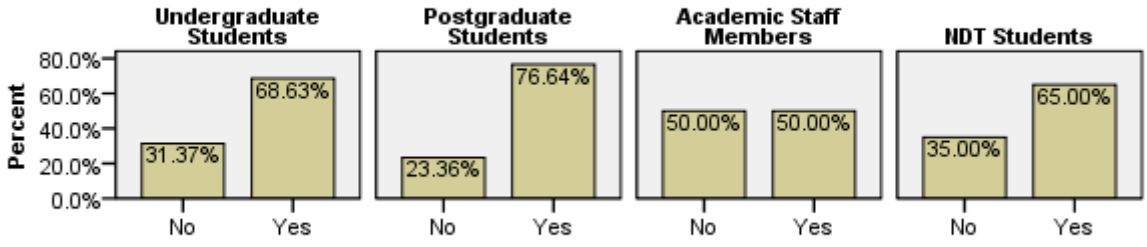


Figure 4.1.3.5: Prior attendance in computer courses by user category

Further, figure 4.1.3.6 depicts whether the respondents' have followed any computer course according to their faculties. It shows that 87.50% users have followed some

computer course in the faculty of Information Technology. It is a very high percentage when compared with others. Although, 60% of the above respondents have followed some computer course in other faculties and institute.

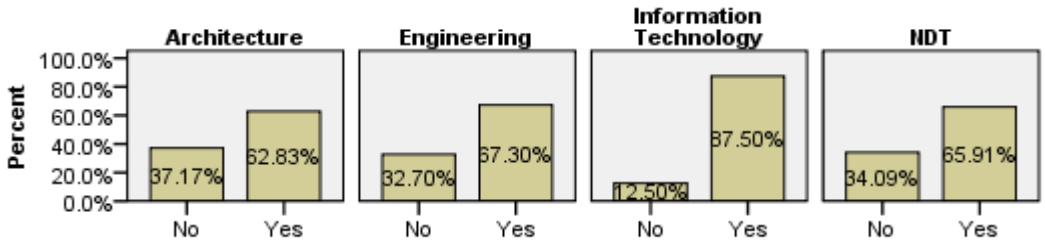


Figure 4.1.3.6: Prior attendance in computer courses by faculty/ institute

Computer knowledge/usage experience (by years) of the respondents is summarized in figure 4.1.3.7 according to their user groups.

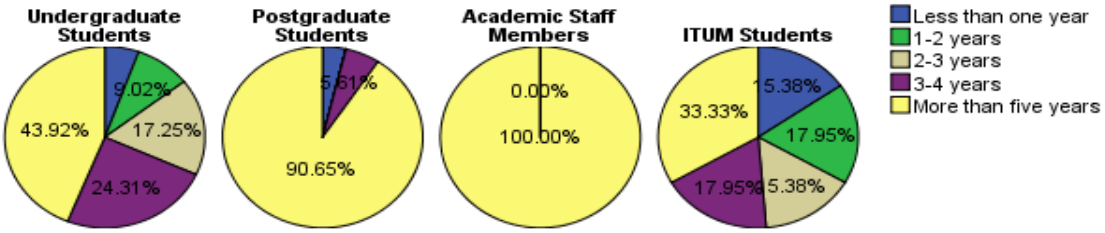


Figure 4.1.3.7: Length of period of computer usage by user category

Figure 4.1.3.7 shows that 100% academic staff members and 90.65% postgraduate students have more than five years experience using computers. But undergraduate and NDT students' computer using experience is lower than the other two groups. This may be because the undergraduate and the NDT students usually come to the university immediately after their A/L examinations. Therefore, they have no chances to use computers for long years.

Further their computer usage experience was examined according to their year of studies in the university. The results depicted in table 4.1.3.2. shows that the undergraduate and NDT students' percentage of computer usage experience gradually increases with the

duration of study. 59.21% of forth year and 75% of fifth year students have more than five years computer usage experience.

Table 4.1.3.2: Length of period of computer used by year of study for undergraduate and NDT students

Year of Study	Usage of Computer					Total
	< one year	1-2 years	2-3 years	3-4 years	5 years <	
Second Year	9	20	18	25	28	100
	9.0%	20.0%	18.0%	25.0%	28.0%	100.0%
Third Year	6	9	24	22	37	98
	6.1%	9.2%	24.5%	22.4%	37.8%	100.0%
Fourth Year	5	0	6	20	45	76
	6.6%	.0%	7.9%	26.3%	59.2%	100.0%
Fifth Year	0	1	2	2	15	20
	.0%	5.0%	10.0%	10.0%	75.0%	100.0%
Total	20	30	50	69	125	294
	6.8%	10.2%	17.0%	23.5%	42.5%	100.0%

The respondents were asked about the Internet connection availability at their homes. The results are summarized according to the user categories in Figure 4.1.3.9.

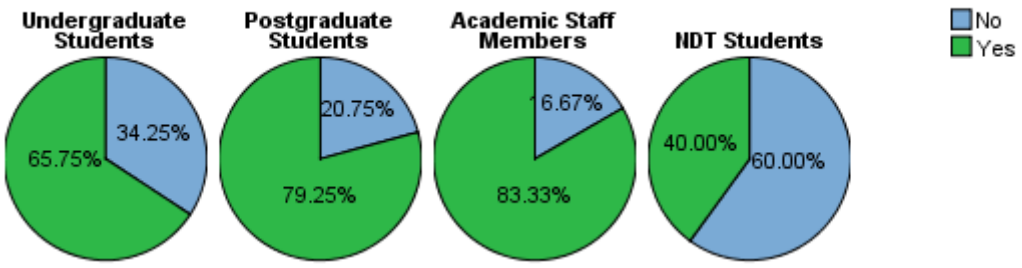


Figure 4.1.3.9: Internet availability at home by user category

Figure 4.1.3.9 shows that, 83.33% academic staff members and 79.25% postgraduate students have internet connection in their homes. But 34.25% undergraduate and 60% NDT students haven't any internet connection in their home. This may be because

academic staff members and the postgraduate students are employed and they have money to implement such extra facilities at their homes and they may believe that they will be useful for their academic activities.

The respondents who have internet connections at their homes were asked the type of the connectivity. Figure 4.1.3.10 describes the results. It clearly shows that 39.52% undergraduate students and 44.19% postgraduate students have broadband connection in their homes. It is comparatively higher than the other connections in these groups. But 48% academic staff members and 31.25% NDT students have ADSL connection in their homes. Within these two groups, ADSL connection users' percentage is higher than the other user groups.

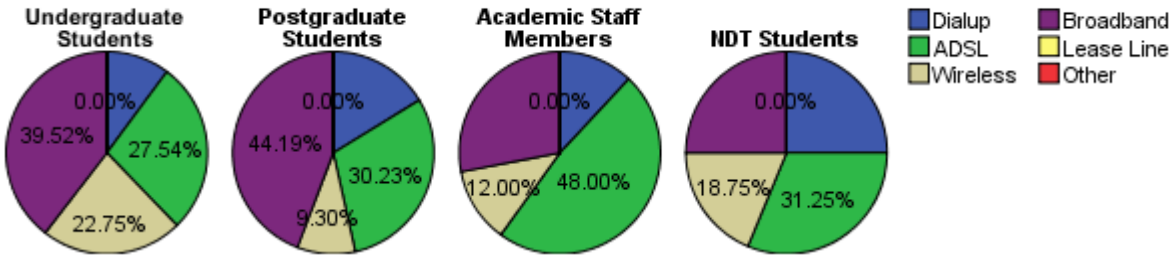
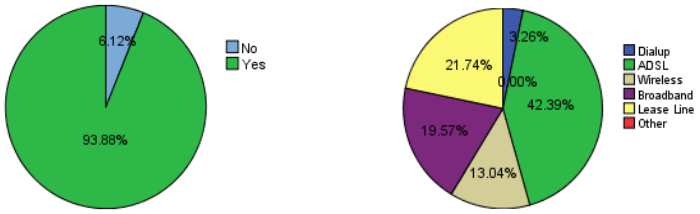


Figure 4.1.3.10: Type of Internet connection at home by user category

Figure 4.1.3.11 further shows that 93.88% postgraduate students' have office Internet connection in offices and 42.39% ADSL connections in offices.



Internet availability at office Type of Internet Connection

Figure 4.1.3.11: Internet availability at office and type of connection for the postgraduate students

4.2 Web searching

The following questions were asked to examine the status of Web searching for study purposes:

- Frequency of using Web & search engines
- Types of search engines use & Web experience
- Purpose of using Web & type of Web information sources used
- Own rating of search skills
- Usage of advanced searching techniques
- Courses or Web searching guides followed on information retrieval

Table 4.2.1: Spearman correlation between Web and computer experience

	Web & Computer experience			
	UG	PG	NDT	All
Correlation Coefficient	.570	.313	.707	.648
Sig. (2-tailed)	.000	.001	.000	.000
N	254	107	39	429

The above table 4.2.1 shows a positive correlation between the Web and computer experience. It is further clarified when the increases in the experience of using computers increases the Web experience.

4.2.1 Frequency of using Web & search engines

Table 4.2.1.1 summarizes the frequency of using Web by user category.

Table 4.2.1.1: Frequency of using Web by category

User Category	Frequency of using Web				Total
	Daily	Weekly	Monthly	Less than once a month	
Undergraduate Students	184 (72.2%)	66 (25.9%)	5 (2.0%)	0 (0.0%)	**255 (100.0%)**
Postgraduate Students	94 (87.9%)	11 (10.3%)	0 (0.0%)	2 (1.9%)	**107 (100.0%)**
Academic Staff Members	**30 (100.0%)**	0 (0.0%)	0 (0.0%)	0 (0.0%)	**30 (100.0%)**
NDT Students	13 (32.5%)	27 (67.5%)	0 (0.0%)	0 (0.0%)	**40 (100.0%)**
Total	321 (74.3%)	104 (24.1%)	5 (1.2%)	2 (0.5%)	**432 (100.0%)**

The most interesting feature of the table 4.2.1.1 is that 100% of respondents are using Web and all academic members are using Web daily. Because all academic staff members have facilities to browse internet in the university and a large percentage (83.33%) of academic staff members have internet connections in their home they can access frequently for searching Web. But comparatively, the daily internet use of NDT students' percentage is very low (32.5%). This may be because most of the students (60%) haven't internet connection their homes.

4.2.1.1 Undergraduate Students

Distribution of the frequency of using Web was further analyzed separately for each user category. Distribution of frequency of using Web of undergraduate students by faculty is shown in figure 4.2.1.1.1.

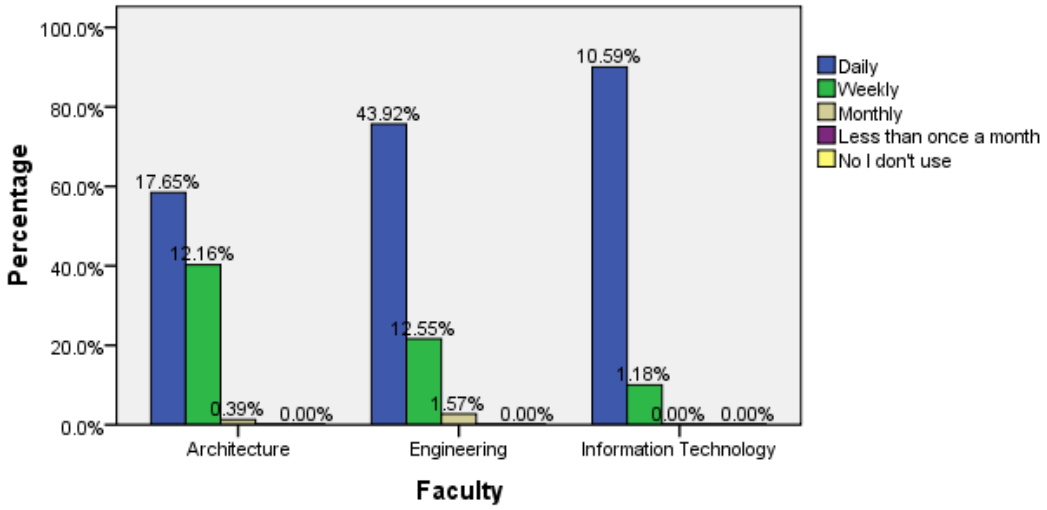


Figure 4.2.1.1.1: Distribution of the frequency of using Web by undergraduate students by faculty

It clearly shows that a higher percentage of undergraduates students in each faculty have used Web on a daily basis. The highest percentage indicated is by students in the faculty of Information Technology. The lowest percentage is shown by the students in the Faculty of Architecture. It is very low compared to the other two faculties. Further, the Kruskal-Wallis One-Way analysis of variance test was performed to test the hypothesis;

H_0: Frequency of using Web is equal for all faculties

versus

H_1: Frequency of using Web is not equal for all faculties

Results of the test were depicted in table 4.2.1.1.1.

Table 4.2.1.1.1: Kruskal-Wallis test result for undergraduate students' frequency in using the Web vs faculties

Ranks

	Faculty	N	Mean Rank
Frequency Web used	Architecture	77	144.91
	Engineering	148	123.86
	Information Technology	30	105.00
	Total	255	

Test Statistics

	Frequency Web used
Chi-Square	12.240
df	2
Asymp. Sig.	.002

According to the above table, the asymptotic significance estimate is 0.002, indicating that the null hypothesis can be rejected at a 0.05 significance level of the alternative hypothesis of at least one difference among the faculty in frequency of Web usage. Further, the mean rank value revealed that the difference is mostly affected by the faculty of information technology. If can be further clarified by figure 4.2.1.1.1 which shows the highest usage by students in the faculty of Information Technology.

The distribution of the frequency of using Web of undergraduate students was further analyzed as their year of studies at the university. The results are summarized in figure 4.2.1.1.2.

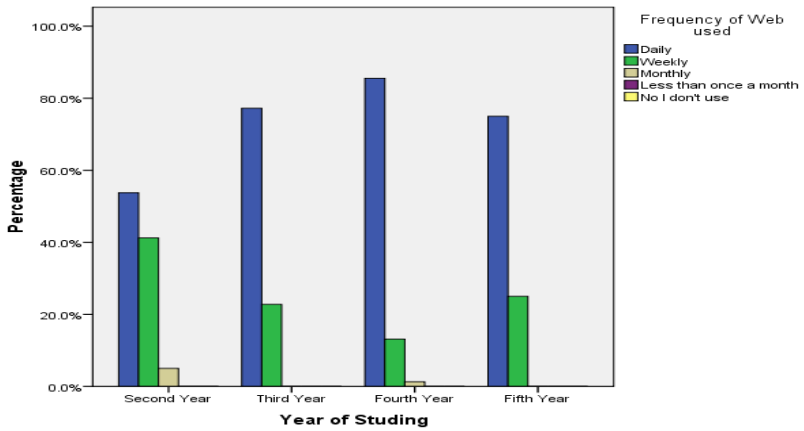


Figure 4.2.1.1.2: Distribution of frequency of using Web by undergraduate students by their year of study

As shown in figure 4.2.1.1.2, when the year of study increases, daily Web usage percentage also increases. But fifth year students daily Web usage percentage is decreasing. All the fifth year students are from the faculty of Architecture where the Web usage is lower. Therefore it may be affected to decrease daily Web using percentage in fifth year students. Further, Chi-square test results revealed that the frequency of using Web significantly differs among the year of study for undergraduate students. The relevant p-value is very small (p=0.000).

The distribution of the frequency of using Web of undergraduate students by their year of study and gender is depicted in Figure 4.2.1.1.3.

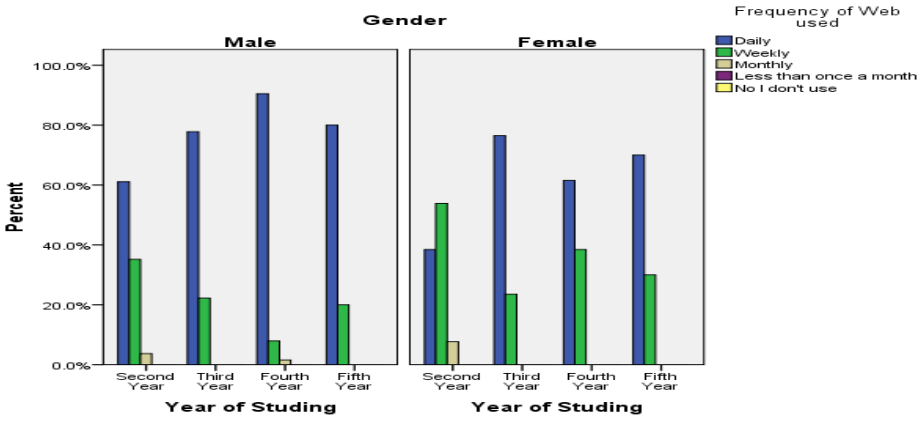


Figure 4.2.1.1.3: Distribution of the frequency of using Web by undergraduate students by their year of study and gender

The above figure shows that male undergraduate students' daily Web using percentage is higher than female students. But third year female students' daily Web using percentage is higher than other years. Further, Chi-square test revealed that the frequency of Web used significantly differs according to the gender of undergraduate students (p-value=0.000).

In addition, the result show a negative correlation between frequency of internet use and internet availability at home (Pearson Correlation, r=-0.335, p-value=0.000). Therefore, it

reveals that if undergraduate students have internet facilities in his/her homes, their internet using frequency could be increased.

4.2.1.2 NDT students

Distribution of frequency of using Web of NDT students by their year of study and gender is shown in figure 4.2.1.2.1.

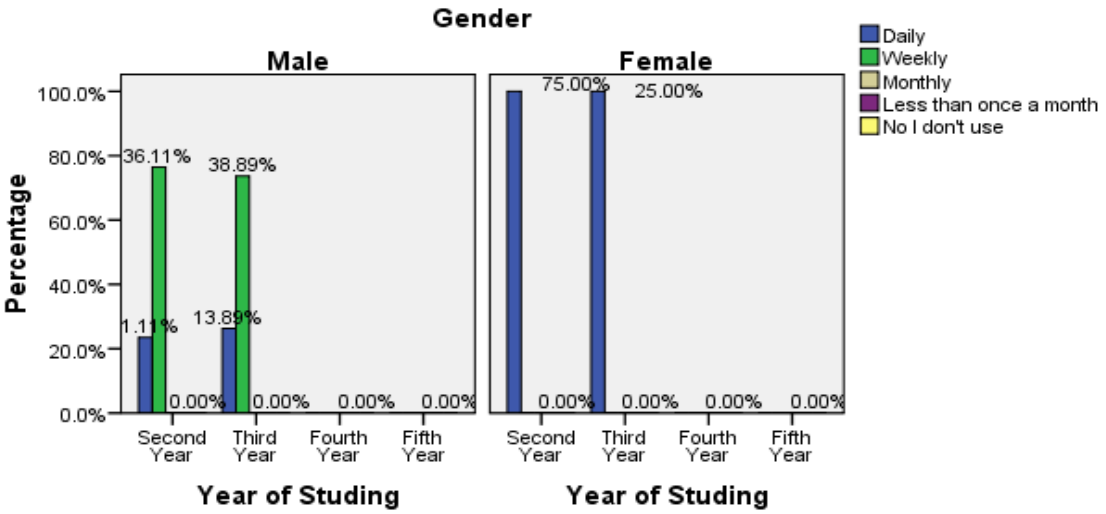


Figure 4.2.1.2.1: Distribution of frequency of using Web by NDT students by their year of study and gender

Figure 4.2.1.2.1 clearly shows that daily Web using percentage is higher for female students than male, and there is no major difference among second and third year female students.

Further, the result show a negative correlation between frequency of internet use and internet availability at home (Pearson Correlation, r=-0.414, p-value=0.008). Therefore, it reveals that if NDT students have internet facility in their homes the internet using frequency will be increased.

4.2.1.3 Postgraduate students

Figure 4.2.1.3.1 describes the distribution of postgraduate students' frequency of using Web by faculty and gender. It shows that male students in the Faculty of Architecture are using Web daily than female students. Other two faculties do not have any considerable difference between male and female postgraduate students.

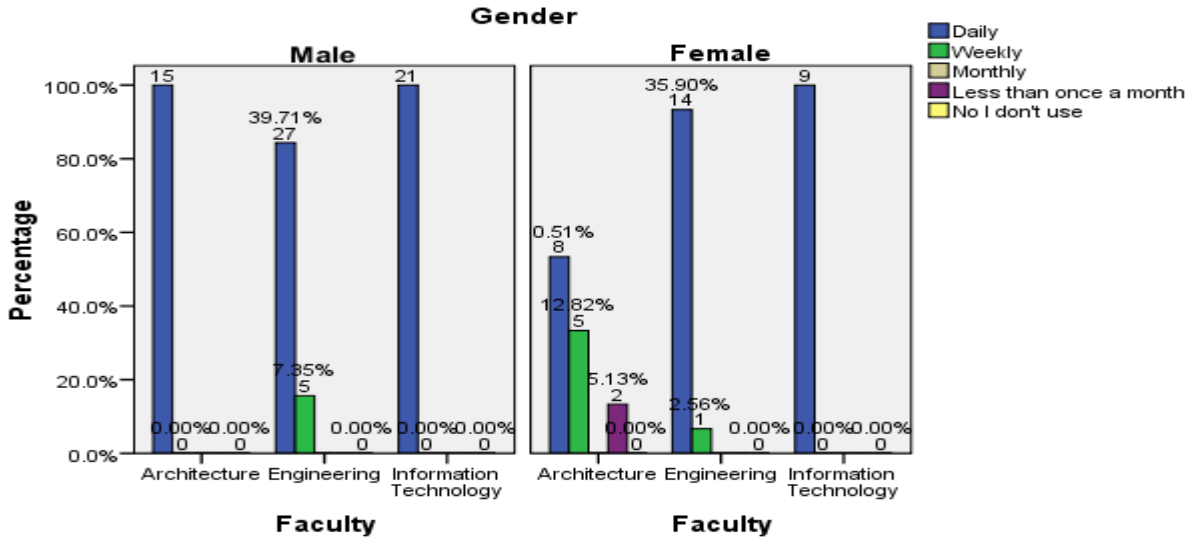


Figure 4.2.1.3.1: Distribution of frequency of using Web by postgraduate students by their faculty and gender

4.2.2 Web using experience

Web using experience of the respondents was analyzed separately for each category.

4.2.2.1 Undergraduate Students

Figure 4.2.2.1.1 summarizes the Web experience of the undergraduate students.

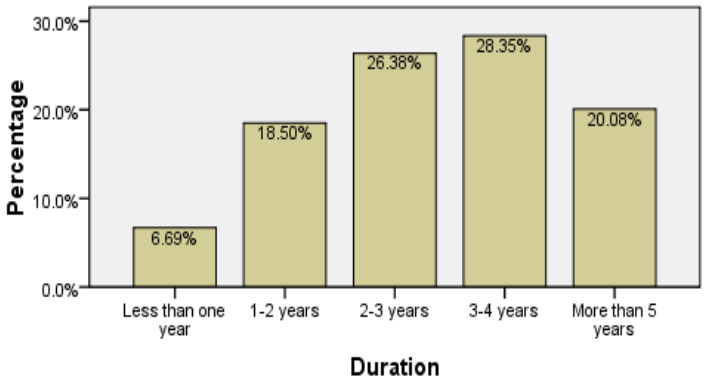


Figure 4.2.2.1.1: Length of period of Web usage for undergraduate students

It shows that 74.86% of undergraduate students have used Web for more than two years. Further, 20.08% of undergraduate students have experience in using Web for more than five years.

Table 4.2.2.1.1 summarizes the experience of Web using according to the year of study for undergraduate students.

Table 4.2.2.1.1: Experience in using the Web by year of study

Year of Studying	Web used experience (Years)					Total
	<1	1-2	2-3	3-4	>5	
Second Year	16 (20.0%)	25 (31.2%)	17 (21.2%)	14 (17.5%)	8 (10.0%)	80 (100.0%)
Third Year	1 (1.3%)	13 (16.7%)	29 (37.2%)	26 (33.3%)	9 (11.5%)	78 (100.0%)
Fourth Year	0 (0.0%)	8 (10.5%)	16 (21.1%)	26 (34.2%)	26 (34.2%)	76 (100.0%)
Fifth Year	0 (0.0%)	1 (5.0%)	5 (25.0%)	6 (30.0%)	8 (40.0%)	20 (100.0%)
Total	17 (6.7%)	47 (18.5%)	67 (26.4%)	72 (28.3%)	51 (20.1%)	254 (100.0%)

The above table shows that when undergraduate students move to the fourth year everyone has more than one year experience for using Web. That means they have a chance to browse the Web during their third year of study.

Further Chi-square test revealed that Web using experience significantly differs among the study of year for undergraduate students (Pearson Chi-Square =71.172, df= 12, p-

value=0.000). It reveals that when students enter the university, the majority do not have any Web experience, but after studying in the university they gradually increase Web browsing.

Table 4.2.2.1.2 summarizes the faculty wise distribution of the Web experience of undergraduate students. It shows that 83.3% percentage of students from the faculty of Information Technology have more than two years Web experience. It is comparatively high when considering the other two faculties. But there is no considerable difference between Engineering and Architecture students.

Table 4.2.2.1.2: Experience in using the Web by undergraduate students by Faculty

Faculty	**Web used experience (Years)**					Total
	<1	**1-2**	**2-3**	**3-4**	**>5**	
Architecture	5 (6.6%)	14 (18.4%)	19 (25.0%)	22 (28.9%)	16 (21.1%)	76 (100.0%)
Engineering	12 (8.1%)	28 (18.9%)	35 (23.6%)	44 (29.7%)	29 (19.6%)	148 (100.0%)
Information Technology	0 (0.0%)	5 (16.7%)	13 (43.3%)	6 (20.0%)	6 (20.0%)	30 (100.0%)
Total	17 (6.7%)	47 (18.5%)	67 (26.4%)	72 (28.3%)	51 (20.1%)	254 (100.0%)

4.2.2.2 NDT Students

The characteristics of the Web using experience of the NDT students are described in this section. Figure 4.2.2.2.1 summarizes the distribution of NDT students' experience in using Web.

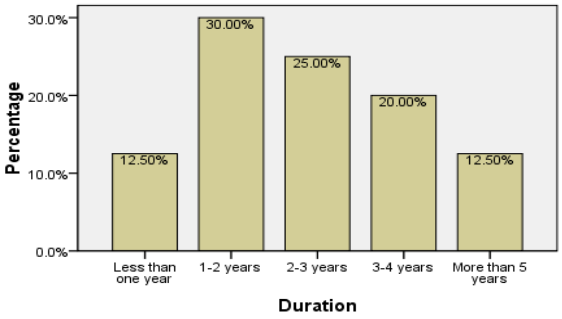


Figure 4.2.2.2.1: Web using experience of NDT students

The above figure shows that the highest (30%) percentage of NDT students have one to two years Web using experience. The lowest (12.5%) percentage of students is in less than one year and more than five years experience groups.

Table 4.2.2.2.1 summarizes NDT students' Web using experience by gender.

Table 4.2.2.2.1: Experience in using the Web by NDT students by gender

Gender	Web used experience (Years)					Total
	<1	1-2	2-3	3-4	>5	
Male	5	12	10	6	3	36
	(13.9%)	(33.3%)	(27.8%)	(16.7%)	(8.3%)	(100.0%)
Female	0	0	0	2	2	4
	(0.0%)	(0.0%)	(0.0%)	(50.0%)	(50.0%)	(100.0%)
Total	5	12	10	8	5	40
	(12.5%)	(30.0%)	(25.0%)	(20.0%)	(12.5%)	(100.0%)

It shows that 33.3% male students have one to two years Web using experience. 100% of female students have more than three years Web experience. But only four female students responded.

4.2.2.3 Postgraduate Students

Figure 4.2.2.3.1 presents the distribution of Web using experience of postgraduate students.

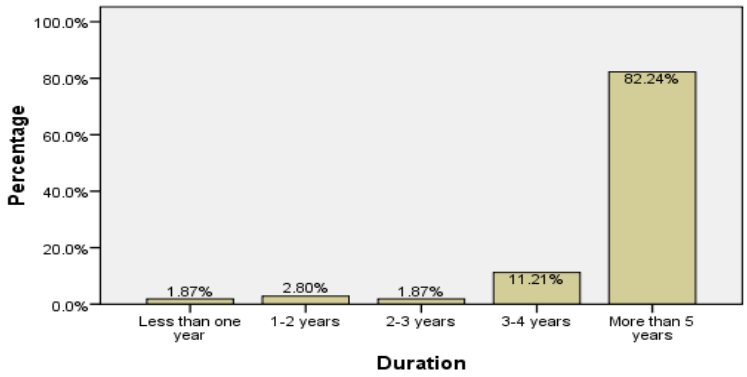


Figure 4.2.2.3.1: Length of period of Web usage for postgraduate students

It clearly shows that most of the postgraduate students (82.24%) have more than five years Web using experience.

Table 4.2.2.3.1 presents the postgraduate students' Web using experience by their respective faculties.

Table 4.2.2.3.1: Web using experience by Faculty for postgraduate students

Faculty	Web used experience (Years)					Total
	<1	1-2	2-3	3-4	>5	
Architecture	2 (6.7%)	2 (6.7%)	0 (0.0%)	3 (10.0%)	23 (76.7%)	30 (100.0%)
Engineering	0 (0.0%)	1 (2.1%)	2 (4.3%)	8 (17.0%)	36 (76.6%)	47 (100.0%)
Information Technology	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (3.3%)	29 (96.7%)	30 (100.0%)
Total	2 (1.9%)	3 (2.8%)	2 (1.9%)	12 (11.2%)	88 (82.2%)	107 (100.0%)

Table 4.2.2.3.1 shows that 96.7% postgraduate students of the faculty of Information Technology have more than five years experience for using Web. Further, Architecture and Engineering postgraduate students have no considerable difference among their Web using experience.

Table 4.2.2.3.2 summarizes that the postgraduate students' Web using experience by their postgraduate courses.

Table 4.2.2.3.2: Web used experience by course for postgraduate students

Course	Web used experience (Years)					Total
	<1	1-2	2-3	3-4	>5	
M.Sc.	2 (2.6%)	2 (2.6%)	1 (1.3%)	6 (7.7%)	67 (85.9%)	78 (100.0%)
M.Eng.	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (10.0%)	9 **(90.0%)**	10 (100.0%)
MBA	0 (0.0%)	1 (5.3%)	1 (5.3%)	5 (26.3%)	12 (63.2%)	19 (100.0%)
Total	2 (1.9%)	3 (2.8%)	2 (1.9%)	12 (11.2%)	88 (82.2%)	107 (100.0%)

Table 4.2.2.3.2 shows that the highest (90%) percentage of Master of Engineering students have more than five years Web experience. But the Web experience of MBA students with more than five years experience is comparatively low; that is 63.2%.

4.2.2.4 Academic staff members

Figure 4.2.2.4.1 presents the Web experience of the teaching faculty members.

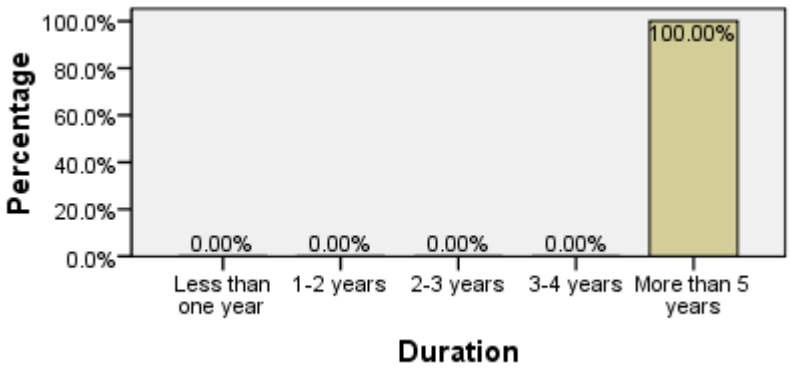


Figure 4.2.2.4.1: Length of period of Web usage for academic staff members

Figure 4.2.2.4.1 depicts that all academic staff members have more than five years experience for using Web. The university has provided unlimited internet facility for academic staff members and most of the academic staff members have Internet connection in their homes. Therefore, their Web experience is higher than other respondents.

4.2.3 Self evaluation of Web searching skills

Own ratings of Web searching skills of the respondents was analyzed separately for their categories under this section.

4.2.3.1 Undergraduate and NDT students

Figure 4.2.3.1.1 summarizes the own ratings of Web searching skills of undergraduate and NDT students. It shows 46.10% students mentioned that she/he has a good knowledge and 2.71 % students mentioned his/her knowledge is poor. Only a 30.85% of students have a satisfactory level of Web searching experiences to manage their Web browsing related activities.

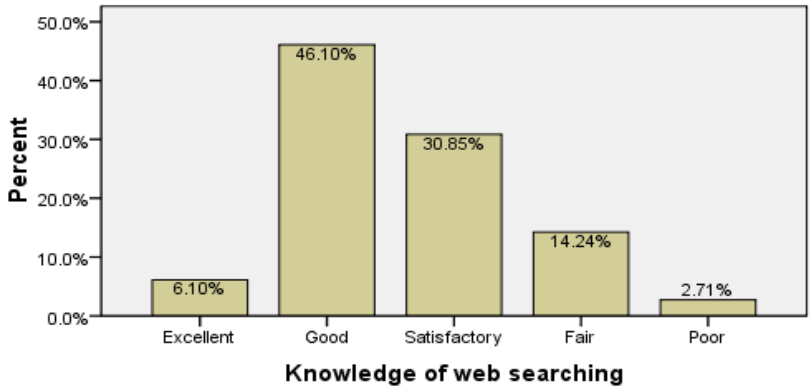


Figure 4.2.3.1.1: Web searching skills of undergraduate and NDT students

Distribution of respondents' Web searching skills by faculty/institute is shown in figure 4.2.3.1.2.

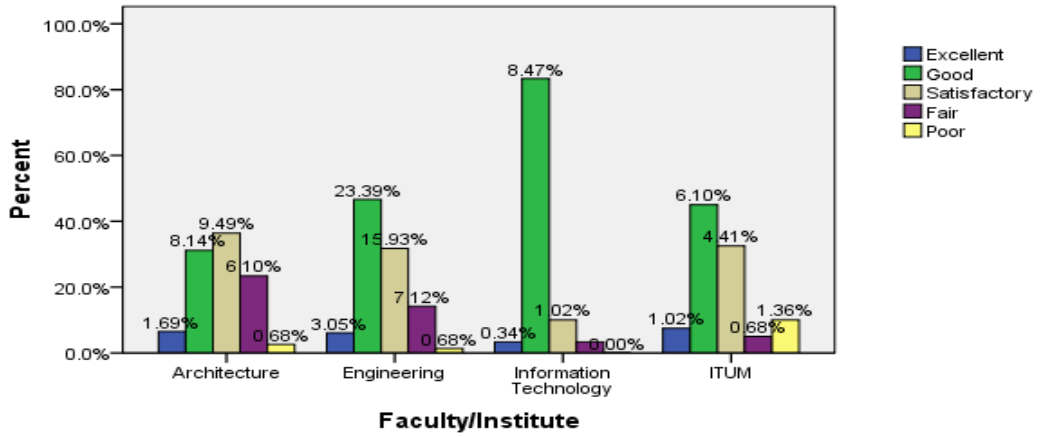


Figure 4.2.3.1.2: Distribution of undergraduate and NDT students' Web searching skills by faculty/institute.

It clearly shows that the majority of the information technology students have mentioned that their Web searching skills are good. Further, the Kruskal-Wallis One-way analysis of variance test was performed to test the hypothesis;

H_0: Web searching knowledge is equal for all faculties/institutes

versus

H_1: Web searching knowledge is not equal for all faculties/institutes

Results of the test are depicted in table 4.2.3.1.1.

Table 4.2.3.1.1: Kruskal-Wallis test results for knowledge of Web searching vs faculty/institute

Ranks

	Faculty/Institute	N	Mean Rank
Web search knowledge	Architecture	77	**170.18**
	Engineering	148	146.17
	Information Technology	30	**101.28**
	NDT	40	147.11
	Total	295	

Test Statistics

	Web search knowledge
Chi-Square	16.415
df	3
Asymp. Sig.	.001

According to the above table, the test statistic has an asymptotic significance estimate of 0.001, indicating that the null hypothesis can be rejected at a 0.05 significant level of the alternative hypothesis of at least one difference among the faculty/institute in knowledge of Web searching. Further, the mean rank revealed that the difference is mostly affected by the students in the faculties of Information Technology and Architecture.

Figure 4.2.3.1.3 presents the distribution of respondents' knowledge of Web searching by year of study.

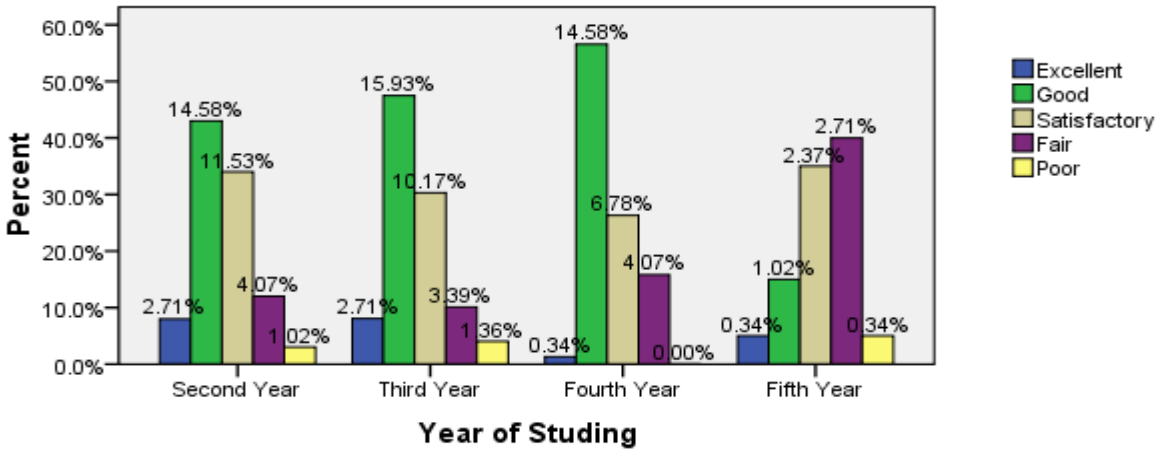


Figure 4.2.3.1.3: Distribution of undergraduate and NDT students' Web searching knowledge by year of study

It clearly shows that the percentage of Web searching knowledge in the level of good or excellent is very low. But all fifth year students are from the Faculty of Architecture. Therefore, this percentage may be very low. Further, Chi-square test revealed that knowledge of Web searching differs significantly among the study of year for the undergraduate and NDT students (Pearson Chi-Square = 25.439, df= 12, p-value=0.013).

4.2.3.2 Postgraduate students

The distribution of postgraduate students' Web searching knowledge by faculty is shown in figure 4.2.3.2.1.

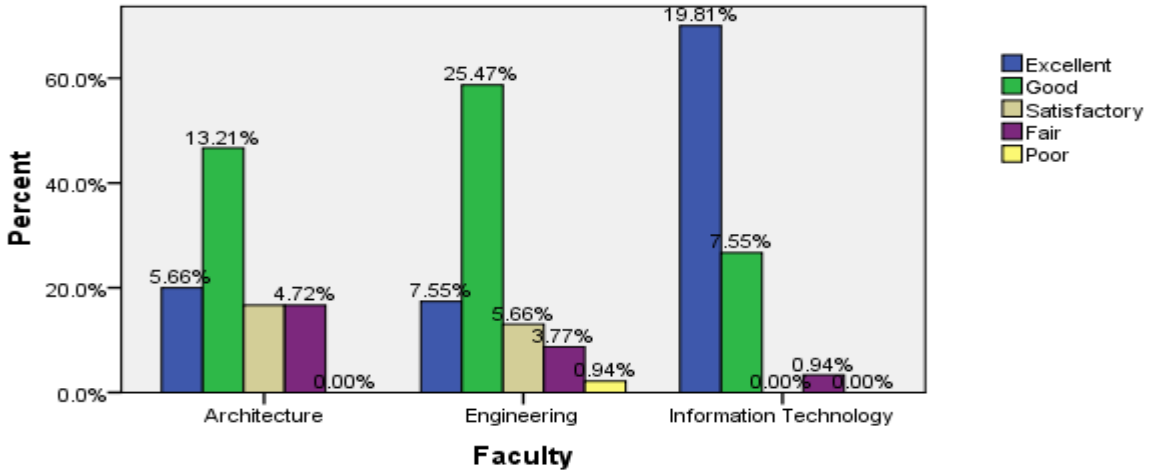


Figure 4.2.3.2.1: Distribution of postgraduate students' Web searching knowledge by faculty

The above figure clearly shows that the majority of postgraduate students in the Faculty of Information Technology mentioned that their Web searching knowledge is excellent. This percentage is very high when compared with the other two faculties.

Therefore, the Kruskal-Wallis One-way analysis of variance test was performed to test the hypothesis;

H_0: Postgraduate students' Web searching knowledge is equal for all the faculties

versus

H_1: Postgraduate students' Web searching knowledge is not equal for all the faculties

Results of the test were depicted in table 4.2.3.2.1.

Table 4.2.3.2.1: Kruskal-Wallis test results for knowledge of Web search vs faculty

Ranks

	Faculty/Institute	N	Mean Rank
Web search knowledge	Architecture	30	63.35
	Engineering	46	61.13
	Information Technology	30	31.95
	Total	106	

Test Statistics

	Web search knowledge
Chi-Square	23.923
df	2
Asymp. Sig.	.000

According to the above table, the test statistic had an asymptotic significance estimate of 0.000, indicating that the null hypothesis can be rejected at a 0.05 significance level of the alternative hypothesis of at least one difference among the faculties in Web searching knowledge. Further, the mean rank revealed that the difference is mostly among the students of Faculty of Information Technology.

Distribution of postgraduate students' Web searching knowledge by age is shown in figure 4.2.3.2.2.

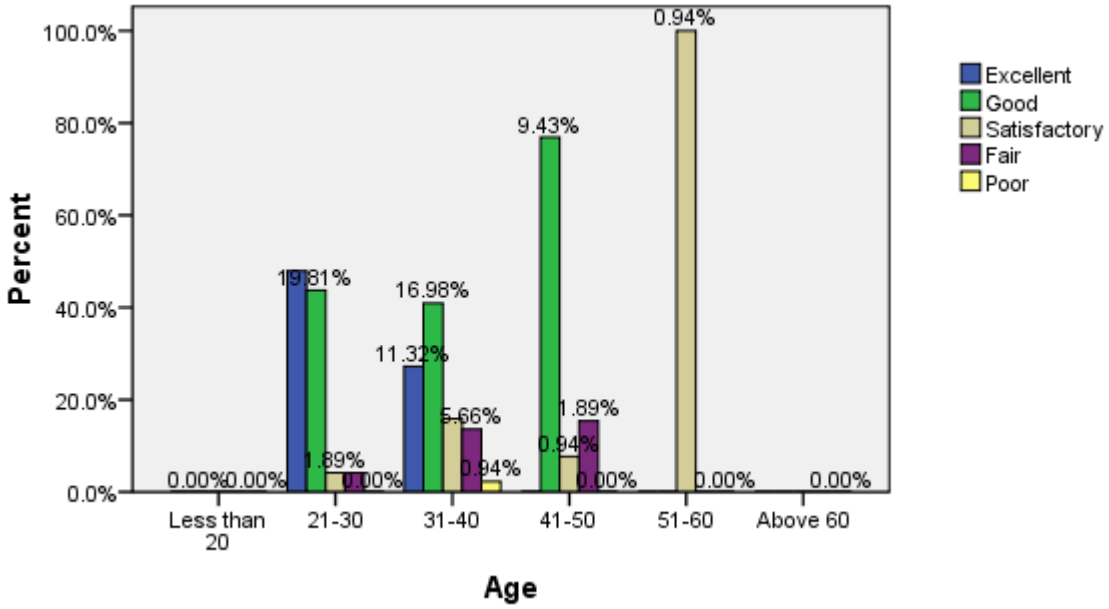


Figure 4.2.3.2.2: Distribution of postgraduate students' Web searching knowledge by age

It clearly shows that majority of post graduate students in the 21-50 years age group mentioned that their Web searching knowledge is good or excellent, but above 50 years students mentioned that their knowledge is in the level of satisfactory. However, only one postgraduate student was in this age group. Further, the highest percentages of postgraduate students in 21-30 years age group mentioned that their Web searching knowledge is excellent.

Further, the Kruskal-Wallis One-way analysis of variance test was performed to test the hypothesis;

H_0: Postgraduate students' Web searching knowledge is equal for all age groups

versus

H_1: Postgraduate students' Web searching knowledge is not equal for all age groups

Results of the test are depicted in table 4.2.3.2.2.

Table 4.2.3.2.2: Kruskal-Wallis test results for knowledge of Web searching vs age

Ranks

	Age	N	Mean Rank
Web search knowledge	21-30	48	42.75
	31-40	44	59.65
	41-50	13	68.35
	Total	105	

Test Statistics

	Web search knowledge
Chi-Square	12.603
df	2
Asymp. Sig.	.002

According to the above table, the test statistic had an asymptotic significance estimate of 0.002, indicating that the null hypothesis can be rejected at a 0.05 significance level of the alternative hypothesis of at least one difference among the age group in Web searching knowledge. Further, the mean rank revealed that the difference is mostly affected by 21-30 years age group.

4.2.3.3 Academic staff members

Figure 4.2.3.3.1 presents the distribution of academic staff members' Web searching knowledge. Figure clearly shows that 33.33% of academic staff members mentioned that their Web searching knowledge is excellent and 50% mentioned that their Web searching knowledge is good.

Figure 4.2.3.3.1: Distribution of academic staff members' Web searching knowledge

Distribution of academic staff members' Web searching knowledge by their faculties is shown in figure 4.2.3.2.2.

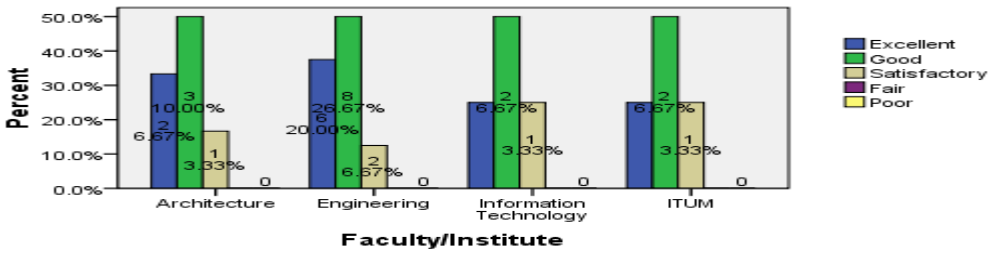


Figure 4.2.3.2.2: Distribution of academic staff members' Web searching knowledge by faculty/institute

It shows that there are no any considerable differences of Web searching knowledge within the faculty/institute. In each faculty most of the staff members mentioned that their Web searching knowledge is good.

4.2.4 Purpose of searching Web

The needs of respondents or the purpose for searching Web is analyzed under this section separately for each user group.

4.2.4.1 Undergraduate and NDT students

Undergraduate and NDT students were asked to rank their purpose of using Web and the mostly used Web information. Results are shown in table 4.2.4.1.1.

Table 4.2.4.1.1: Rank percentages of Web using purpose of undergraduate and NDT students

Rank / Purpose	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	Total	%
Research	43 (20.5%)	42 (20.0%)	39 (18.6%)	27 (12.9%)	52 (24.8%)	5 (2.4%)	2 (1.0%)	- -	- -	- -	210 (100%)	71.2
Education	70 (27.7%)	59 (23.3%)	44 (17.4%)	55 (21.7%)	16 (6.3%)	8 (3.2%)	-	1 (0.4%)	-	-	253 (100%)	85.8
Entertainment	53 (22.7%)	62 (26.2%)	78 (33.5%)	27 (11.6%)	12 (5.2%)	1 (0.4%)	-	-	-	-	233 (100%)	79.0
E-Business	2 (6.9%)	-	-	1 (3.4%)	4 (13.8%)	10 (34.5%)	2 (6.9%)	1 (3.4%)	-	9 (31%)	29 (100%)	9.8
Send or receive email	**111 (50%)**	21 (9.5%)	49 (22.1%)	29 (13.1%)	11 (5.0%)	1 (0.5%)	-	-	-	-	222 (100%)	75.3
Acquiring information of personal interest	2 (2.6%)	10 (13.2%)	8 (10.5%)	8 (10.5%)	6 (7.9%)	19 (25%)	16 (21.1%)	2 (2.6%)	5 (6.6%)	-	76 (100%)	25.8
Downloading software	11 (7.3%)	49 (32.7%)	14 (9.3%)	44 (29.3%)	22 (14.7%)	9 (6.0%)	1 (0.7%)	-	-	-	150 (100%)	50.8
Communication	3 (4.0%)	5 (6.7%)	7 (9.3%)	13 (17.3%)	6 (8.0%)	15 (20.0%)	14 (18.7%)	12 (16.0%)	-	-	75 (100%)	25.4
Electronic publishing	-	-	-	-	2 (9.5%)	6 (28.6%)	1 (4.8%)	2 (9.5%)	9 (42.9%)	1 (4.8%)	21 (100%)	7.1
Career development	-	-	-	-	3 (11.1%)	2 (7.4%)	7 (25.9%)	7 (25.9%)	7 (25.9%)	1 (3.7%)	27 (100%)	9.2

Table 4.2.4.1.1 shows that majority of the undergraduate and NDT students (85.8%) use Web for education purposes. 79% and 75.3% of students use Web for entertainment and to send or receive e-mail respectively. As highlighted in the table, the 50% of undergraduate and NDT students have ranked their 1st preference to send or receive e-mails. Thus, they mostly used Web to send or receive e-mails. Further, 71.2% and 50.8% of students used Web for research and downloading software respectively. Few number of students used Web for other purposes.

Friedman test was employed to see the difference between purposes of using Web under an adjustment. Since all the respondents have not used the rank value for all mention purposes, the least preference rank was replaced as the missing values in order to have similar number or responses. The test hypothesis is:

H_0: there is no difference among ranks versus H_1: there is any difference among ranks

The result of the Friedman test is depicted in table 4.2.4.1.2.

Table 4.2.4.1.2: Friedman test results for the significance of purposes order of using Web

Ranks

Purpose	Mean Rank
Research	4.28
Education	3.26
Entertainment	3.47
E-Business	7.87
Send or receive email	3.33
Acquiring information of personal interest	7.09
Downloading software	5.36
Communication	7.21
Electronic publishing	7.98
Career development	7.97
Other	8.18

Test Statistics[a]

N	295
Chi-Square	1.621E3
df	10
Asymp. Sig.	.000

a. Friedman Test

According to the resultant table 4.2.4.1.2 the asymptotic significance estimate is very small, indicating that the null hypothesis can be rejected at a 0.05 significance level. Therefore it says that there is any significant difference among ranks.

Table 4.2.4.1.3 shows the Friedman test results for the purposes of using Web by undergraduate and NDT students according to their faculty/institute. It clearly shows that each faculty/institute significance estimate for the test was very small (0.000) and was less than 0.05. Therefore, there is a significant difference among faculty/institute on the purpose of searching Web. Further, Architecture and Engineering students' first

preference is given for using Web to send or receive e-mails. But IT and NDT students' first preference is given for education purposes.

Table 4.2.4.1.3: Friedman test results for the significance of purposes of using Web by faculty/institute

Purpose	Faculty (Mean Rank)			
	Architecture	**Engineering**	**IT**	**NDT**
Research / Project work	3.62	4.14	5.40	5.26
Education	3.46	3.29	**2.73**	**3.15**
Entertainment	3.69	3.15	3.73	4.02
E-Business	7.81	7.92	7.75	7.89
Send or receive email	**3.38**	**2.99**	4.25	3.76
Acquiring information of personal interest	6.91	7.21	7.12	6.99
Downloading software	5.99	5.15	4.88	5.32
Communication	7.35	7.47	6.65	6.35
Electronic publishing	7.94	8.11	7.78	7.74
Career development	7.84	8.12	7.85	7.74
Other	8.01	8.45	7.85	7.78
Test Statistics				
N	77	148	30	40
Chi-Square	423.738	899.869	148.604	181.608
df	10	10	10	10
Asymp. Sig.	0.000	0.000	0.000	0.000

Friedman test results performed to see the differences between gender and among the year of study of undergraduate students are summarized in table 4.2.4.1.4. It clearly shows that the asymptotic significance estimate for the test was very small (0.000) for each category and it was less than 0.05. Therefore, there is some difference among ranks. Further, second year and fifth year students are mostly using Web for education purposes, and third year students are mostly using Web to send or receive e-mail, and forth year students are using Web for their entertainment. But when we consider the gender difference, it shows that most of male undergraduate students are using Web to send or receive e-mails and female students' first preference is given for education purposes.

Table 4.2.4.1.4: Friedman test results for the significance of purposes of using Web by year of study and gender

Mean Ranks						
Purpose	**Year of study**				**Gender**	
	2nd	**3rd**	**4th**	**5th**	**Male**	**Female**
Research	4.61	4.32	3.78	4.38	4.61	3.51
Education	**3.25**	3.35	3.28	**2.82**	3.25	**3.29**
Entertainment	3.38	3.79	**2.97**	4.18	3.25	3.98
E-Business	7.74	8.02	7.83	7.90	7.88	7.84
Send or receive email	3.54	**3.34**	2.98	3.50	**3.19**	3.64
Acquiring information of personal interest	7.10	6.78	7.49	7.12	7.17	6.91
Downloading software	5.61	5.52	4.88	5.22	5.14	5.90
Communication	7.19	7.06	7.43	7.18	7.26	7.07
Electronic publishing	7.77	7.95	8.33	7.90	7.96	8.03
Career development	7.86	7.76	8.39	7.90	4.61	7.75
Other	7.96	8.12	8.64	7.90	3.25	8.09
Test Statistics						
N	100	99	76	20	208	87
Chi-Square	529.093	501.223	493.715	111.273	1.180E3	461.132
df	10	10	10	10	10	10
Asymp. Sig.	0.000	0.000	0.000	0.000	.000	.000

4.2.4.2 Postgraduate students

Table 4.2.4.2.1 shows that Friedman test results for the purpose of using Web by postgraduate students.

Table 4.2.4.2.1: Friedman test results for the significance of purposes of using Web for postgraduate students

Ranks

Purpose	Mean Rank
Research / Project work	3.00
Education	3.23
Entertainment	5.47
E-Business	8.03
Send or receive email	2.26
Acquiring information of personal interest	6.87
Downloading software	5.14
Communication	7.09
Electronic publishing	8.23
Career development	7.86
Other	8.82

Test Statistics[a]

N	107
Chi-Square	652.938
df	10
Asymp. Sig.	0.000

a. Friedman Test

It shows that the asymptotic significance estimate for the test was very small (0.000) and it was less than 0.05. Therefore, there is some difference among ranks. Further, the most of the postgraduate students use Web to send or receive e-mail mostly and secondly, they use Web to find information for research / project work.

Table 4.2.4.2.2 summarizes the Friedman test results for the purpose of using Web fore post graduate students by course, gender and marital status.

Table 4.2.4.2.2: Friedman test results for the significance of purposes of using Web for the postgraduate students by course, gender and marital status

Mean Ranks							
Purpose	**Course**			**Gender**		**Marital Status**	
	M.Sc	**M.Eng**	**MBA**	**Male**	**Female**	**Married**	**Unmarried**
Research	2.67	3.60	4.03	3.46	2.18	2.96	3.01
Education	3.15	3.65	3.34	3.10	3.45	3.16	3.29
Entertainment	5.54	5.45	5.21	5.39	5.62	5.30	5.71
E-Business	8.11	8.40	7.50	7.85	8.35	7.80	8.35
Send or receive email	**2.40**	**1.10**	**2.29**	**2.38**	**2.06**	**2.52**	**1.91**
Acquiring information of personal interest	6.67	8.10	7.08	6.75	7.09	7.41	5.98
Downloading software	5.36	5.05	4.26	4.81	5.71	5.73	4.37
Communication	7.08	6.70	7.37	7.02	7.22	7.24	6.91
Electronic publishing	8.15	8.60	8.34	8.38	7.97	8.03	8.48
Career development	8.04	6.40	7.92	7.94	7.73	7.48	8.49
Other	8.84	8.95	8.66	8.93	8.63	8.36	9.50
Test Statistics							
N	78	10	19	68	39	64	41
Chi-Square	488.817	68.276	109.595	396.102	267.805	378.963	276.868
df	10	10	10	10	10	10	10
Asymp. Sig.	0.000	0.000	0.000	0.000	0.000	0.000	0.000

It clearly shows that, within the each category there is a significance difference in the purpose of searching Web. For all categories, first preference is given to send or receive e-mails. M.Sc. and M.Eng. students' second preference is for research and MBA students' second preference is for education purposes. Further, male post graduate students' second preference is for education purposes and females' second preference is

for research activities. However, it shows that most of the postgraduate students are using Web to send or receive e-mails, research activities and education purposes.

4.2.4.3 Academic staff members

Table 4.2.4.3.1 shows that the Friedman test results for the purposes of using Web by academic staff members.

Table 4.2.4.3.1: Friedman test results for the significance of purposes of using Web for the academic staff members

Ranks

purpose	Mean Rank
Research	1.82
Education	4.73
Entertainment	6.18
E-Business	8.05
Send or receive email	2.40
Acquiring information of personal interest	5.77
Downloading software	6.77
Communication	6.87
Electronic publishing	7.92
Career development	7.52
Other	7.98

Test Statistics

N	30
Chi-Square	178.767
df	10
Asymp. Sig.	**0.000**

Test statistics show a significant difference among each purpose of using Web by academic staff members. Mean rank value shows academic staff members' main purpose of using web is to find information for their research activities. Secondly, the purpose of using Web is to send or receives e-mails and thirdly, for education. Therefore, it shows that most of the academic staff members using Web for research activities and to send or receive e-mail and for education purposes.

Table 4.2.4.3.2 presents the Friedman test results for the purposes of using Web by academic staff members according to their academic position. It shows an availability of a

significant difference among the purposes of Web searching by senior lecturers and lecturer probationary. Further, the mean rank values shows senior lecturers primarily use Web to find information for research purposes and their second priority is given to send or receive e-mails. But probationary lecturers' first priority is given to send or receive e-mails and second priority is given to find information for research activities.

Table 4.2.4.3.2: Friedman test results for the significance of purposes of using Web by academic staff members according to their academic positions

Ranks		
Purpose	**Sen. Lecture**	**Lecture(Prob.)**
Research	1.76	1.80
Education	4.85	5.10
Entertainment	6.24	5.00
E-Business	8.09	7.70
Send or receive email	2.76	1.30
Acquiring information of personal interest	5.57	7.20
Downloading software	6.85	6.60
Communication	7.04	6.10
Electronic publishing	7.76	8.40
Career development	7.24	8.40
Other	7.85	8.40
Test Statistics		
N	23	5
Chi-Square	131.543	36.759
df	10	10
Asymp. Sig.	**0.000**	**0.000**

4.2.5. Web resources/sources used

Usage of Web re/sources by respondents were analyzed separately for user categories.

4.2.5.1 Undergraduate and NDT students

Table 4.2.5.1.1 presents the types of Web resources used by undergraduate and NDT students. It shows that 91.5% undergraduate and NDT students are using Web

information resources except their categories and 86.1% of students are using sources in the same manner.

Table 4.2.5.1.1: Case summary of undergraduate and NDT students for Web information resources and sources used

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Resources	270	91.5%	25	8.5%	295	100.0%
Sources	254	86.1%	41	13.9%	295	100.0%

Table 4.2.5.1.2 presents the number of students and their percentages of types of Web resources used.

Table 4.2.5.1.2: Frequencies of undergraduate and NDT students Web information resources usage

		Responses		Percent of Cases
		N	Percent	
Resources	News papers	133	21.8%	49.3%
	Journals/Magazines	130	21.3%	48.1%
	E-Books	182	29.8%	67.4%
	Reference resources	163	26.7%	60.4%
	Other	3	.5%	1.1%
Total		611	100.0%	226.3%

It clearly shows that most of the undergraduate and NDT students (60.4%) have used e-books and considerable percentage (60.4%) of students have used reference resources (e.g. encyclopedia, dictionaries).

Table 4.2.5.1.3 depicts the percentages of the undergraduate and NDT students' usage of Web sources. It clearly shows most of the students (70.1%) are using social networks. Other resources using percentage is low.

Table 4.2.5.1.3: Frequencies of undergraduate and NDT students Web information sources usage

		Responses		Percent of Cases
		N	Percent	
Sources	News portal	66	11.0%	26.0%
	Web portal	59	9.8%	23.2%
	Library catalog	84	14.0%	33.1%
	Database	89	14.9%	35.0%
	Social networks	178	29.7%	70.1%
	Subject gateways	46	7.7%	18.1%
	Blogs	77	12.9%	30.3%
Total		599	100.0%	235.8%

Table 4.2.5.1.4 summarizes the percentages of undergraduate and NDT students' usage of Web resources by their types.

Table 4.2.5.1.4: Frequencies of undergraduate and NDT students Web information resources usage by faculty/institute

Resources	Faculty/Institute				Total
	Architecture	Engineering	Information Technology	NDT	
News papers	30 (42.3%)	63 (47.4%)	17 (56.7%)	23 (63.9%)	**133**
Journals/Magazines	39 (54.9%)	66 (49.6%)	14 (46.7%)	11 (30.6%)	**130**
E-Books	43 (60.6%)	95 (71.4%)	29 **(96.7%)**	15 (41.7%)	**182**
Reference resources	35 (49.3%)	86 (64.7%)	23 (76.7%)	19 (52.8%)	**163**
Other	0 (0.0%)	0 (0.0%)	3 (10.0%)	0 (0.0%)	**3**
Total	**71**	**133**	**30**	**36**	**270**

It shows that the majority (96.7%) of students in the faculty of Information Technology are using e-books. Further, considerable percentage of Architecture (60.6%) and Engineering (71.4%) students are using e-books. But, the majority (63.9%) of NDT students are using news papers. It reveals that a comparatively high percentage of NDT students are using news papers, but the percentage of other resources usage is comparatively low.

Table 4.2.5.1.5 presents undergraduate and NDT students' Web sources usage by faculty/institute.

Table 4.2.5.1.5: Frequencies of undergraduate and NDT students Web information sources usage by faculty/institute

Sources	Faculty/Institute				Total
	Architecture	Engineering	IT	NDT	
News portal	11 17.7%	34 26.6%	12 41.4%	9 25.7%	**66**
Web portal	10 16.1%	26 20.3%	17 58.6%	6 17.1%	**59**
Library catalog	19 30.6%	42 32.8%	10 34.5%	13 37.1%	**84**
Database	24 38.7%	44 34.4%	11 37.9%	10 28.6%	**89**
Social networks	47 **75.8%**	87 **68.0%**	25 **86.2%**	19 **54.3%**	**178**
Subject gateways	11 17.7%	23 18.0%	8 27.6%	4 11.4%	**46**
Blogs	11 17.7%	42 32.8%	19 65.5%	5 14.3%	**77**
Total	**62**	**128**	**29**	**35**	**254**

It also shows that the highest percentages of students in each faculty/institute are using social networks. But NDT students' percentage is low for those using social networks, comparative with other three faculties.

4.2.5.2 Postgraduate students

Table 4.2.5.2.1 shows that 98.1% and 93.5% postgraduate students are using Web resources and other sources respectively.

Table 4.2.5.2.1: Frequencies of postgraduate students Web information resources and sources usage

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Resources	105	98.1%	2	1.9%	107	100.0%
Sources	100	93.5%	7	6.5%	107	100.0%

Table 4.2.5.2.2 presents the usage of Web resources by postgraduate students.

Table 4.2.5.2.2: Frequencies of postgraduate students Web information resources usage

		Responses		Percent of Cases
		N	Percent	
Resources	News papers	71	23.1%	67.6%
	Journals/Magazines	91	29.6%	**86.7%**
	E-Books	79	25.7%	75.2%
	Reference resources	66	21.5%	62.9%
Total		**307**	**100.0%**	**292.4%**

It shows that 86.7% of postgraduate students have used online journals/magazines. This percentage is comparatively high when compared with other resources. The second highest percentage (75.2%) is for e-books and a considerable percentage of postgraduate students are using news papers (67.6%) and reference resources (62.9%) too.

Table 4.2.5.2.3 presents the frequencies of Web information resources used by postgraduate students according to their faculties and degree courses.

Table 4.2.5.2.3: Frequencies of postgraduate students Web information resources usage by faculty and course

Resources	Faculty/Institute			Course			Total
	Architecture	Engineering	IT	M.Sc	M.Eng	MBA	
News papers	15	34	22	50	8	13	**71**
	53.6%	72.3%	73.3%	65.8%	80.0%	68.4%	
Journals/ Magazines	25	38	28	65	9	17	**91**
	89.3%	**80.9%**	93.3%	**85.5%**	**90.0%**	**89.5%**	
E-Books	16	33	30	58	8	13	**79**
	57.1%	70.2%	**100.0%**	76.3%	80.0%	68.4%	
Reference resources	10	28	28	47	5	14	**66**
	35.7%	59.6%	93.3%	61.8%	50.0%	73.7%	
Total	**28**	**47**	**30**	**76**	**10**	**19**	

It shows that most of the post graduate students in Architecture and Engineering faculties use journals/magazines. The highest percentages are shown in table 4.2.5.2.3 for journals/magazine use for all faculties, except the Faculty of Information Technology.

100% of postgraduate students are using e-books and 93.3% are using journals/magazine. The percentage of using reference resources (35.7%) by postgraduate students in the Faculty of Architecture is comparatively low.

Table 4.2.5.2.4 presents the frequency of Web sources usage for postgraduate students.

Table 4.2.5.2.4: Frequencies of postgraduate students Web information sources usage

		Responses		Percent of Cases
		N	Percent	
Sources	News portal	50	16.1%	50.0%
	Web portal	51	16.4%	51.0%
	Library catalog	32	10.3%	32.0%
	Database	52	16.7%	52.0%
	Social networks	62	19.9%	**62.0%**
	Subject gateways	28	9.0%	28.0%
	Blogs	35	11.3%	35.0%
	Other	1	.3%	1.0%
Total		**311**	**100.0%**	**311.0%**

It shows that the highest percentage (62%) of postgraduate students use social networks. A considerable percentage of postgraduate students use news portal (50%), Web portal (51%) and databases (52%).

Table 4.2.5.2.5 summarizes postgraduate students' frequency of Web information sources usage by faculty and courses.

It shows that 51.7% Architecture students use databases; 65.1% of Engineering students use social networks; 82.1% of IT students use Web portals; 64.4% of M.Sc students use social networks; 66.7% of M.Eng students use news portals and 61.1% of MBA students use Web portals. But the highest percentage (75%) of Blogs are use d by students in the Faculty of Information Technology. This percentage is comparatively very high when compared with the other two faculties.

Table 4.2.5.2.5: Frequencies of postgraduate students Web information sources usage by faculty and course

Sources	Faculty/Institute			Course			Total
	Architecture	Engineering	IT	M.Sc	M.Eng	MBA	
News portal	12 (41.4%)	20 (46.5%)	18 (64.3%)	37 (50.7%)	6 **(66.7%)**	7 (38.9%)	**50**
Web portal	7 (24.1%)	21 (48.8%)	23 **(82.1%)**	39 (53.4%)	1 (11.1%)	11 **(61.1%)**	**51**
Library catalog	9 (31.0%)	13 (30.2%)	10 (35.7%)	24 (32.9%)	3 (33.3%)	5 (27.8%)	**32**
Database	15 **(51.7%)**	19 (44.2%)	18 (64.3%)	44 (60.3%)	4 (44.4%)	4 (22.2%)	**52**
Social networks	14 (48.3%)	28 **(65.1%)**	20 (71.4%)	47 **(64.4%)**	5 (55.6%)	10 (55.6%)	**62**
Subject gateways	5 (17.2%)	13 (30.2%)	10 (35.7%)	18 (24.7%)	4 (44.4%)	6 (33.3%)	**28**
Blogs	3 (10.3%)	11 (25.6%)	21 (75.0%)	28 (38.4%)	0 (0.0%)	7 (38.9%)	**35**
Other	0 (0.0%)	1 (2.3%)	0 (0.0%)	1 (1.4%)	0 (0.0%)	0 (0.0%)	**1**
Total	**29**	**43**	**28**	**73**	**9**	**18**	

4.2.5.3 Academic staff members

Table 4.2.5.3.1 presents the percentages of using Web sources and resources by academic staff members.

Table 4.2.5.3.1: Frequencies of academic staff members Web information resources and sources usage

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Resources	30	100.0%	0	.0%	30	100.0%
Sources	28	93.3%	2	6.7%	30	100.0%

It shows that 100% academic staff members use Web resources and 93.3% use Web sources. Two academic staff members didn't mention which kinds of Web resources they have used.

Table 4.2.5.3.2 summarizes the frequency of using Web information resources by academic staff members.

Table 4.2.5.3.2: Frequencies of academic staff members Web information resources usage

		Responses		Percent of Cases
		N	Percent	
Resources	News papers	18	19.8%	60.0%
	Journals/Magazines	27	29.7%	**90.0%**
	E-Books	23	25.3%	76.7%
	Reference resources	23	25.3%	76.7%
	Total	**91**	**100.0%**	**303.3%**

It shows that the highest percentage (90%) of academic staff members use journal/magazine. Considerable percentage of academic staff members use e-books (76.7%), reference resources (76.7%) and news papers (60%), but a comparatively low percentages of staff members use news papers.

Table 4.2.5.3.3 summarizes the frequencies of Web information sources usage of academic staff members.

Table 4.2.5.3.3: Frequencies of academic staff members Web information sources usage

		Responses		Percent of Cases
		N	Percent	
Sources	News portal	13	12.4%	46.4%
	Web portal	14	13.3%	50.0%
	Library catalog	15	14.3%	53.6%
	Database	25	23.8%	**89.3%**
	Social networks	15	14.3%	53.6%
	Subject gateways	9	8.6%	32.1%
	Blogs	12	11.4%	42.9%
	Other	2	1.9%	7.1%
	Total	**105**	**100.0%**	**375.0%**

It shows that databases usage (89.3%) of academic members is higher than others. More than 50% of academic staff members used Library catalog (53.6%), Social networks

(53.6%) and Web portal (50%). But the percentage of their usage of subject gateways is low (32.1%).

Table 4.2.5.3.4 presents the Web resources usage of academic staff members.

Table 4.2.5.3.4: Frequencies of academic staff members Web information resources usage by faculty and academic position

Resources	Faculty/Institute				Academic Position			Total
	Architecture	**Engineering**	**IT**	**NDT**	**Professor**	**Senior Lecturer**	**Lecturer (Prob.)**	
News papers	6	9	0	3	2	13	3	18
	100.0%	56.2%	.0%	75.0%	**100.0%**	56.5%	60.0%	
Journals/ Magazines	5	16	2	4	2	20	5	27
	83.3%	**100.0%**	50.0%	**100.0%**	**100.0%**	**87.0%**	**100.0%**	
E-Books	4	13	4	2	2	17	4	23
	66.7%	81.2%	**100.0%**	50.0%	**100.0%**	73.9%	80.0%	
Reference resources	5	14	2	2	1	18	4	23
	83.3%	87.5%	50.0%	50.0%	50.0%	78.3%	80.0%	
	6	16	4	4	2	23	5	

Table shows that 100% of Architecture academic staff members use news papers; 100% of Engineering faculty staff members use journals/magazine and 100% of IT faculty staff members use e-books. All professors use news papers, journals/magazines and e-books. But there were only two professors in the sample. 87% senior lecturers and 100% probationary lectures use journals/magazine.

Table 4.2.5.3.5 summarizes the frequencies of using Web information sources by academic staff members according to the faculty and academic position.

It shows 100% of IT and NDT; 87.5% of Engineering and 83.3% Architecture academic staff members use databases. Although, 100% of staff members from faculty of IT use Web portal, social networks and subject gateways.

Table 4.2.5.3.5: Frequencies of academic staff members Web information sources usage by faculty and academic position

Sources	Faculty/Institute				Academic Position			Total
	Architecture	Engineering	IT	NDT	Professor	Senior Lecturer	Lecturer (Prob.)	
News portal	3	8	1	1	2	9	2	13
	50.0%	50.0%	50.0%	25.0%	**100.0%**	42.9%	40.0%	
Web portal	3	9	2	0	1	10	3	14
	50.0%	56.2%	**100.0%**	.0%	50.0%	47.6%	60.0%	
Library catalog	3	10	1	1	1	11	3	15
	50.0%	62.5%	50.0%	25.0%	50.0%	52.4%	60.0%	
Database	5	14	2	4	1	19	5	25
	83.3%	**87.5%**	**100.0%**	**100.0%**	50.0%	**90.5%**	**100.0%**	
Social networks	4	8	2	1	1	10	4	15
	66.7%	50.0%	**100.0%**	25.0%	50.0%	47.6%	80.0%	
Subject gateways	2	4	2	1	1	6	2	9
	33.3%	25.0%	**100.0%**	25.0%	50.0%	28.6%	40.0%	
Blogs	3	8	1	0	0	10	2	12
	50.0%	50.0%	50.0%	.0%	.0%	47.6%	40.0%	
Other	0	2	0	0	0	2	0	2
	.0%	12.5%	.0%	.0%	.0%	9.5%	.0%	
Total	6	16	2	4	2	21	5	

4.2.6 Frequency of using search engine

Factors related to the search engine usage of respondents are described under this section.

4.2.6.1 Undergraduate and NDT students

To determine the usage of search engines, respondents were asked to rate the frequency of their usage. Figure 4.2.6.1.1 depicts the results of undergraduate and NDT students. As shown in the figure, the frequency of search engine usage differs widely. About 67.01%

use the search engine daily; 27.89% use once a week; 3.4% use monthly and 1.7% use less than once a month.

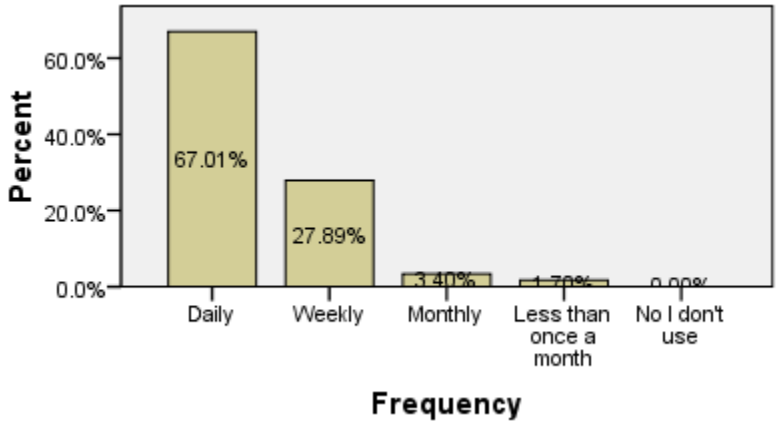


Figure 4.2.6.1.1: Frequency of using search engines

Figure 4.1.4.6.2 presents the frequency of using Web search engines by faculty/institute.

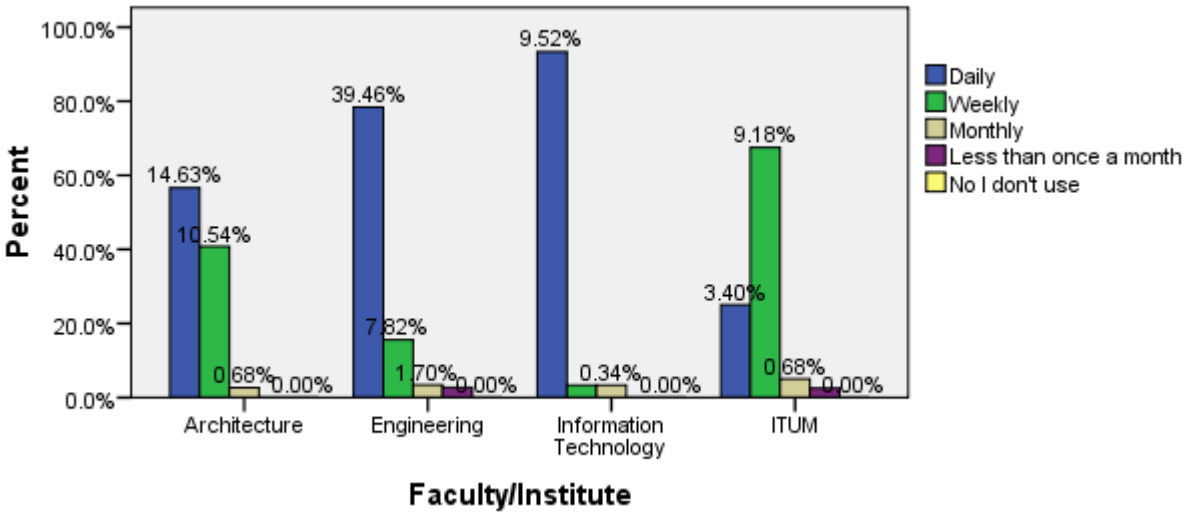


Figure 4.2.6.1.2: Frequency of using search engines by faculty/institute

It clearly shows that the percentage of NDT students using search engine daily is very low (less than 30%) and about 90% of IT students use search engine daily.

Table 4.2.6.1.1 shows the Kruskal-Wallis test results for frequency of using search engines by faculty/Institute. It revealed that frequency of Web search engine usage has

statistically significant difference according to the faculty/institute. Further, it revealed that the most significance difference is between Architecture and NDT students.

Table 4.2.6.1.1: Kruskal-Wallis test results for Frequency of search engines usage by faculty/Institute for the undergraduate and NDT students

Ranks

	Faculty/Institute	N	Mean Rank
Frequency of use search engines	Architecture	77	**162.69**
	Engineering	148	137.12
	IT	30	127.45
	NDT	40	**175.39**
	Total	295	

Test Statistics

	used search engines
Chi-Square	14.100
df	3
Asymp. Sig.	0.003

4.2.6.2 Postgraduate students

Figure 4.2.6.2.1 summarizes the postgraduate students' frequency of using search engines according to their faculties.

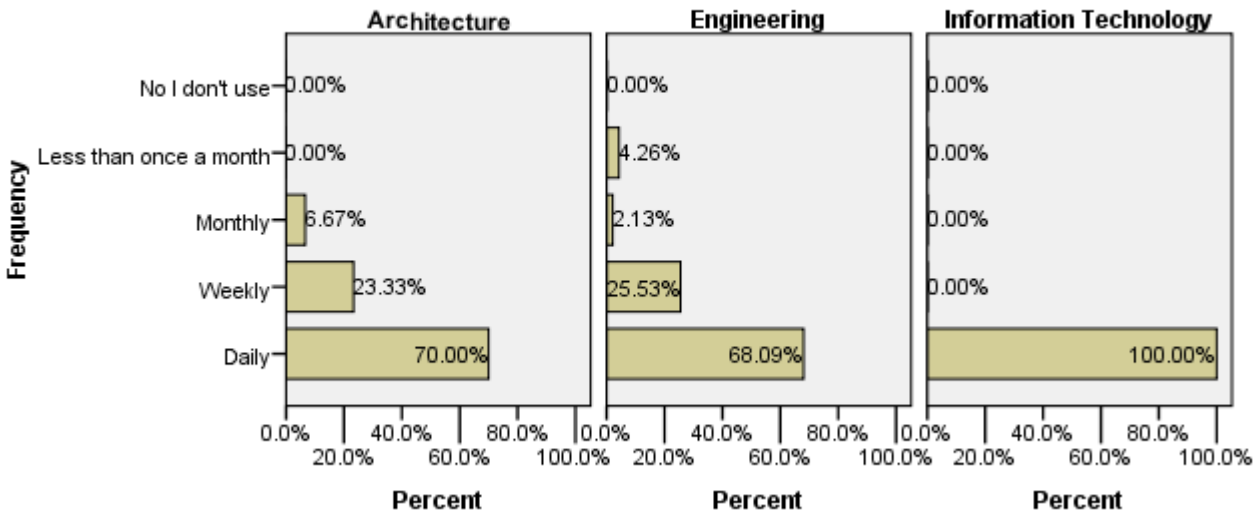


Figure 4.2.6.2.1: Frequency of using search engines by postgraduate students by faculty

Figure shows that the majority of the postgraduate students use search engine daily. Further, 100% of postgraduate students in the IT faculty use search engines daily. Further table 4.2.6.2.1 depicted the Kruskal-Wallis test results.

Table 4.2.6.2.1: Kruskal-Wallis test results for Frequency of search engines usage of postgraduate students as faculties

Ranks

	Faculty/Institute	N	Mean Rank
Frequency of Search engines usage	Architecture	30	58.03
	Engineering	47	59.09
	IT	30	**42.00**
	Total	107	

Test Statistics

	Search engines usage
Chi-Square	11.854
df	2
Asymp. Sig.	**0.003**

It clearly shows that the frequency of search engine usage is statistically significant with faculty and most difference are shown in the Faculty of IT.

Figure 4.2.6.2.2 summarizes the frequency of Web search engine usage of postgraduate students by gender.

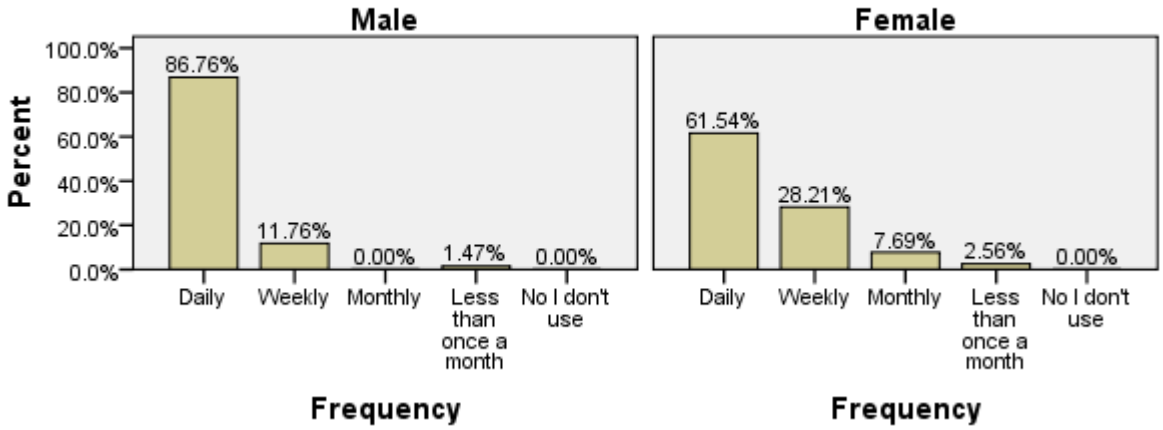


Figure 4.2.6.2.2: Frequency of search engines usage by gender for postgraduate students

It clearly shows that the percentage of male postgraduate students (86.76%) using search engines daily is higher than the percentage of female students (61.54%). Further table

4.1.4.6.3 revealed a significant difference between males and females using search engines.

Table 4.2.6.2.2: Kruskal-Wallis test results for Frequency of search engines usage by gender for the postgraduate students

Ranks

	Gender	N	Mean Rank
Frequency of search engines usage	Male	68	48.95
	Female	39	**62.81**
	Total	107	

Test Statistics

	Freq. of search engines usage
Chi-Square	9.369
df	1
Asymp. Sig.	**0.002**

Table 4.2.6.2.3 presents the frequency of search engine usage in each age group of postgraduate students.

Table 4.2.6.2.3: Frequency of search engines usage by age for postgraduate students

		Frequency				
		Daily	Weekly	Monthly	Less than once a month	Total
Age	21-30	44	4	0	0	48
		91.7%	8.3%	.0%	.0%	100.0%
	31-40	29	11	3	1	44
		65.9%	25.0%	6.8%	2.3%	100.0%
	41-50	10	2	0	1	13
		76.9%	15.4%	.0%	7.7%	100.0%
	51-60	0	2	0	0	2
		.0%	100.0%	.0%	.0%	100.0%
Total		83	19	3	2	107
		77.6%	17.8%	2.8%	1.9%	100.0%

It shows that the percentage of using search engine daily is high for postgraduate students in 21-30 years age group. Further, 51-60 years age group students' daily usage is very low (0%). But in that group, only two students were included. Therefore, the Kruskal-

Wallis test was performed ignoring these two students. The Kruskal-Wallis test results are shown in table 4.2.6.2.3.

Table 4.2.6.2.3: Kruskal-Wallis test results for frequency of search engines usage by postgraduate students according to their age

Ranks

	Age	N	Mean Rank
Search engines usage	21-30	48	**46.17**
	31-40	44	60.01
	41-50	13	54.50
	Total	105	

Test Statistics

	Search engines usage
Chi-Square	9.525
df	2
Asymp. Sig.	**0.009**

Table 4.2.6.2.3 revealed a significant difference among age groups under 0.05. The most significant age group is 21-30 years. It means that the frequency of Web search engine usage is high among young post graduate students.

4.2.6.3 Academic staff members

Figure 4.2.6.3.1 presents the frequency of Web search engine usage of academic staff members.

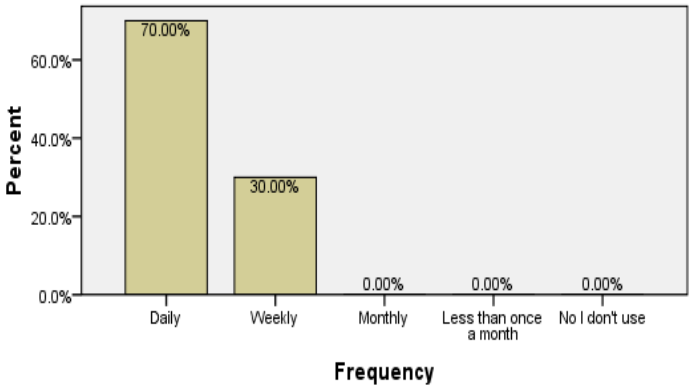


Figure 4.2.6.3.1: Frequency of search engines usage by academic staff members

It shows that 70% of academic staff members use Web search engines daily and 30% of members use them weekly.

4.2.7 Search engines usage

To determine the often used search engines, respondents were asked to rate the search engines used by them. Figure 4.2.7.1 summarizes the results.

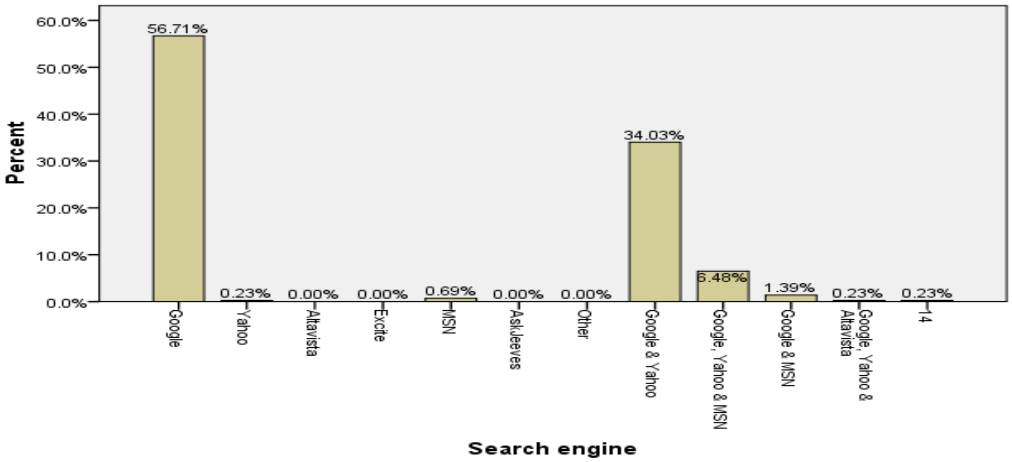


Figure 4.2.7.1: Mostly used search engines

It shows that, 56.71% respondents use only Google; 34.03% use Google and Yahoo; 6.48% uses Google, Yahoo and MSN; 1.39% use Google and MSN. It revealed that 98.61% (56.71+34.03+6.48+1.39) of respondents use Google. This may be because Google is a well known and mostly used search engine.

4.2.8 Experience of using search engines

Table 4.2.8.1 presents the Pearson correlation coefficients and significance values for Web using experience and experience of search engine usage. It shows a positive linear correlation between Web experience and search engine experience. It implies that an increase in Web usage results in an increase in search engine usage.

Table 4.2.8.1: Correlations between Web experience and search engine experience

		Web used experience	Experience of using search engines
Web used experience	Pearson Correlation	1	.836
	Sig. (2-tailed)		.000
	N	431	430
Experience of using search engines	Pearson Correlation	.836	1
	Sig. (2-tailed)	.000	
	N	430	431

Table 4.2.8.2 presents the correlations between Web experience and search engine experience for undergraduates, NDT students and postgraduate students. It also shows that each user category has a significant positive correlation.

Table 4.2.8.2: Correlations between Web experience and search engine experience for undergraduate and NDT students and postgraduate students

		Web used experience		Experience of using search engines	
		Ug&NDT	Pg	Ug&NDT	Pg
Web used experience	Pearson Correlation	1	1	.849	.579
	Sig. (2-tailed)			.000	.000
	N	294	107	293	107
Experience of using search engines	Pearson Correlation	.849	.579	1	1
	Sig. (2-tailed)	.000	.000		
	N	293	107	294	107

4.2.8.1 Undergraduate and NDT students

Figure 4.2.8.1.1 summarizes the experience of search engine usage for undergraduate and NDT students. It shows the highest percentage of students in 3-4 years group and above 67% of students have more than two years search engine using experience.

Figure 4.2.8.1.2 depicts the search engine usage of undergraduate and NDT students according to their year of study.

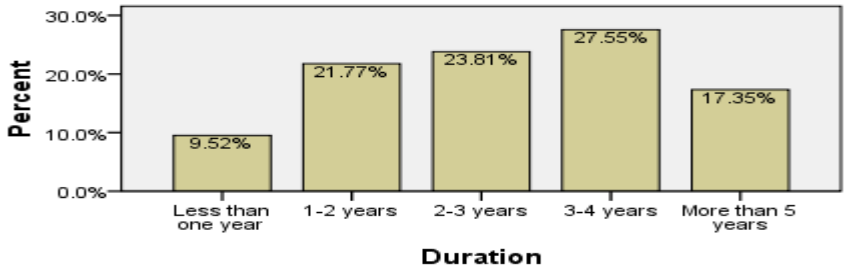


Figure 4.2.8.1.1: Length of search engine usage for the undergraduate and NDT students

Figure 4.2.8.1.2 reveals that when increasing the year of study, length of search engine usage also increases. But fourth to fifth year percentage of search engine experience in 1-2 years has also increased. Because all fifth year students are Architecture and comparatively their Web search engine experience is low.

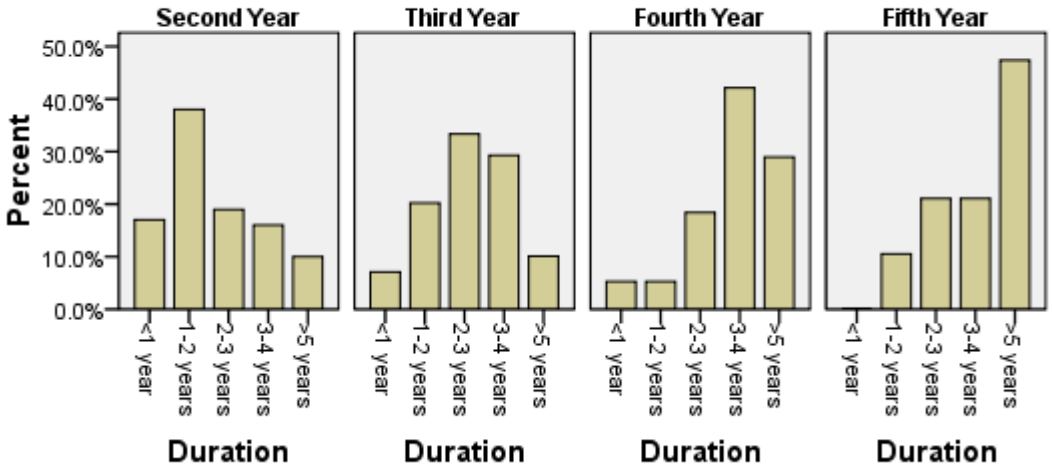


Figure 4.2.8.1.2: Length of search engine usage by undergraduate and NDT students according to year of study

4.2.8.2 Postgraduate students

Figure 4.2.8.2.1 depicts the search engine usage experience of postgraduate students and it shows that 78.5% postgraduate students have more than five years experience in using Web search engines. Further, there isn't anyone who has less than one year search engine using experience.

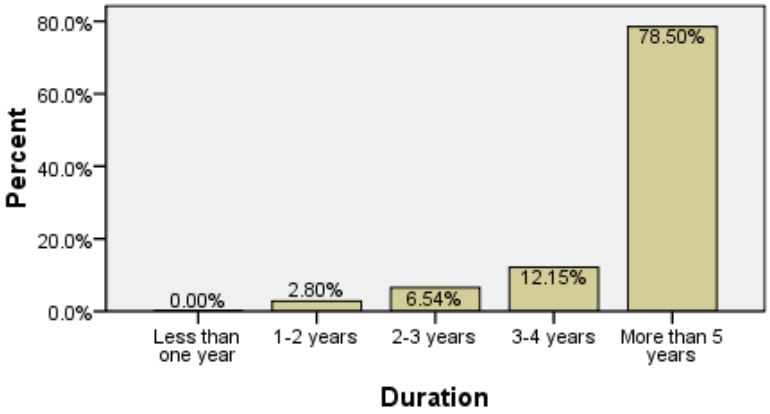


Figure 4.2.8.2.1: Search engine using experience of postgraduate students

Figure 4.2.8.2.2 presents the search engine using experience of postgraduate students by faculty.

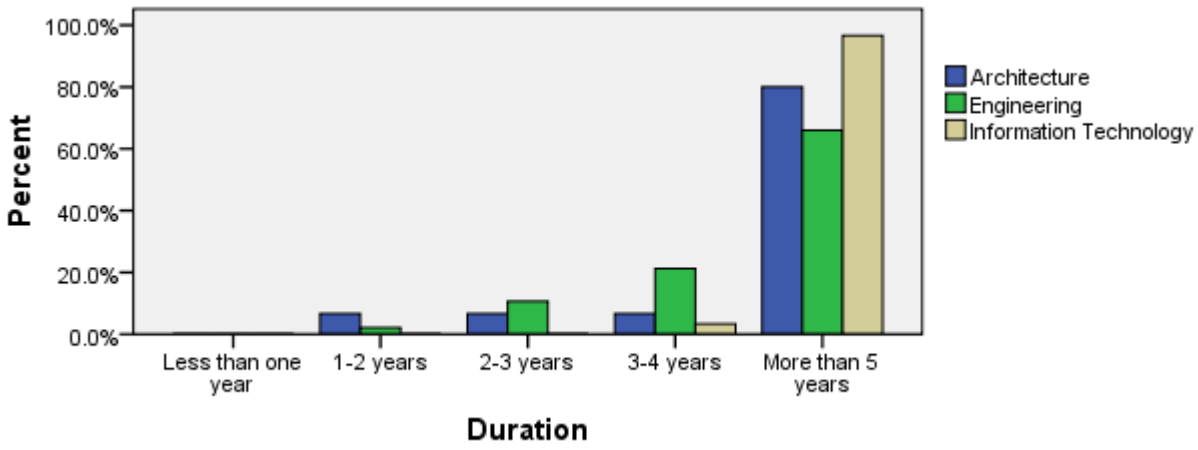


Figure 4.2.8.2.2: Search engine using experience of postgraduate students by faculty

It shows that the post graduate students in the Faculty of IT have a higher experience by using search engines. It is more than five years. Further, table 5.2.8.2.1 presents the Kruskal-Wallis test results. Table shows that a significant difference among faculties and Engineering students' experience of Web search engines has mostly affected the difference.

Table 4.2.8.2.1: Kruskal-Wallis test results for length of search engines usage by faculty for postgraduate students

Ranks

	Faculty	N	Mean Rank
Experience of using search engines	Architecture	30	54.13
	Engineering	47	**47.61**
	Information Technology	30	63.88
	Total	107	

Test Statistics

	Experience of using search engines
Chi-Square	9.800
df	2
Asymp. Sig.	**0.007**

Figure 4.2.8.2.3 presents the length of search engine usage for postgraduate students by the courses followed.

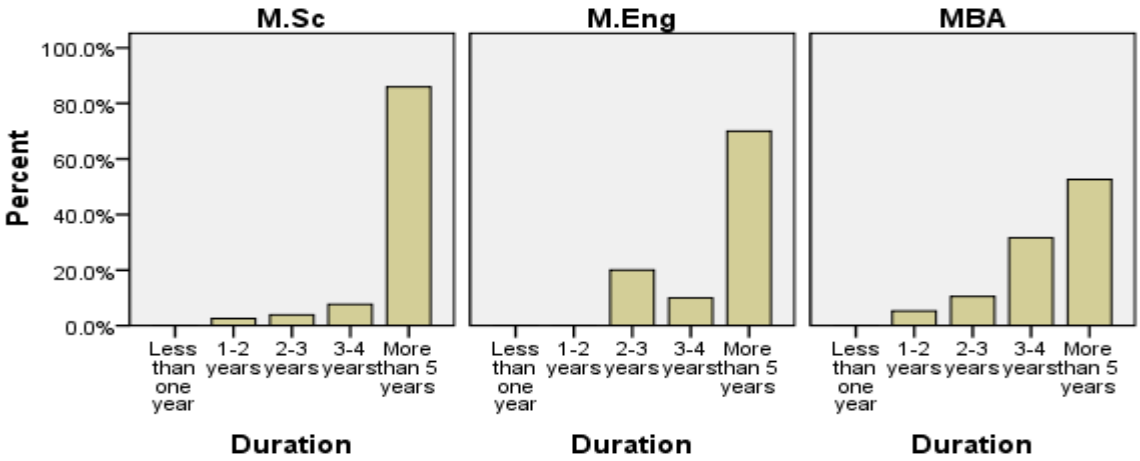


Figure 4.2.8.2.3: Length of search engine usage for the postgraduate students by course

It shows that the highest percentage with more than five years experience students is in the M.Sc. course group. In addition, comparatively MBA students' Web search engine experience is lower than others. Further, Kruskal-Wallis test results of table 4.2.8.2.2 reveals that a significant difference among the courses and MBA students' experience of Web search engine has mostly affected.

Table 4.2.8.2.2: Kruskal-Wallis test results for length of search engines usage by course for postgraduate students

Ranks

	Course	N	Mean Rank
Experience of using search engines	M.Sc	78	57.89
	M.Eng	10	48.95
	MBA	19	**40.68**
	Total	107	

Test Statistics

	Experience of using search engines
Chi-Square	9.705
df	2
Asymp. Sig.	**0.008**

4.2.8.3 Academic staff members

Figure 4.2.8.3.1presents the search engine using experience of academic staff members.

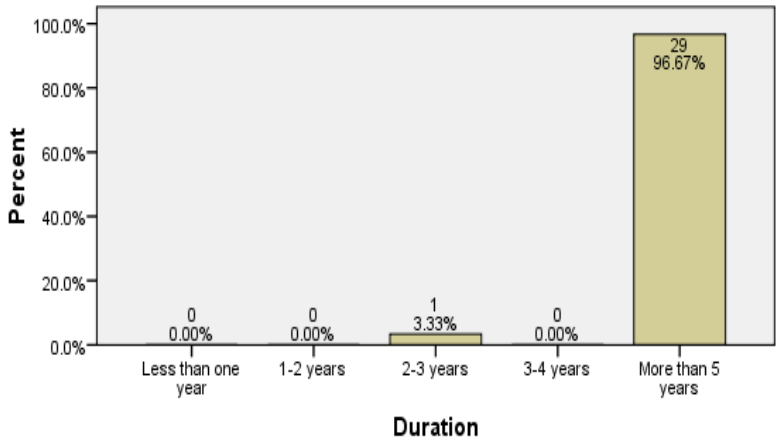


Figure 4.2.8.3.1: Length of search engine usage for academic staff members

The above figure shows that 96.67% of academic staff members have more than five years experience in using search engines. In addition, only one staff member has 2-3 years experience working in the faculty of Architecture.

4.2.9 Participation of formal training on Web information searching

Figure 4.2.9.1 presents the details of participation of formal training on Web information searching for each user category.

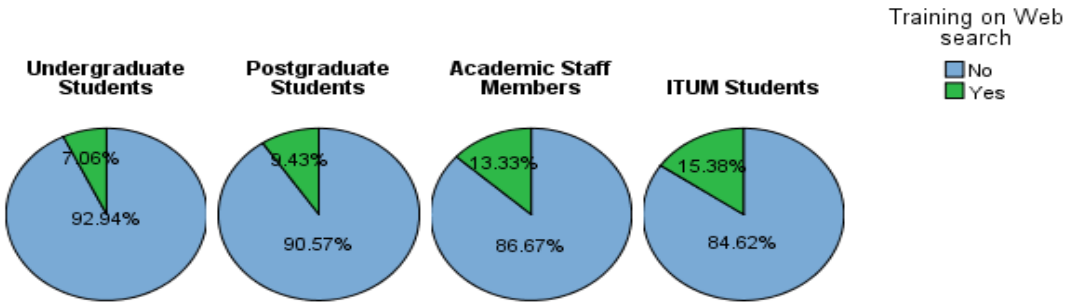


Figure 4.2.9.1: Participation in formal training on Web information searching by user category

It shows that the most of respondents have not participated in any formal training on Web information searching.

4.2.10 Use of Web search guides

Figure 4.2.10.1 depicts the information on using Web search guides by respondents.

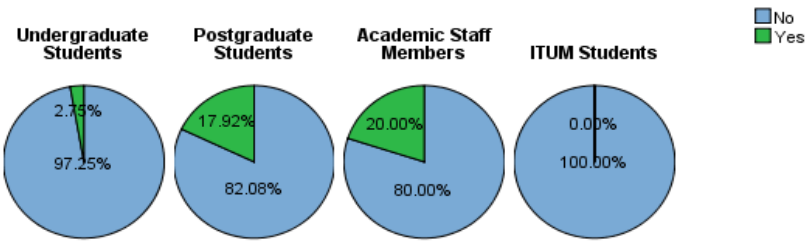


Figure 4.2.10.1: Usage of Web search guides on Web information searching

4.2.11 Using Web searching strategies

When Web users have formulated search queries using mind maps or subject maps concepts, free text searching, word strings, full phrase search, advance Boolean operators, word stems, term truncation or wildcards methods. Therefore, participants were asked to indicate if they had used the above mentioned methods. Results are summarized in table 4.2.11.1. It presents the usage of each search strategies for each user category and Pearson Chi-Square results. Pearson Chi-Square results shows whether or not there exists a significance difference among the user categories.

Table 4.2.11.1: Web searching strategies and its usage for each user group and Pearson Chi-Square test statistics

Searching strategies	User category								Pearson Chi-Square		
	UG Students		PG Students		Acc. Staff		NDT Students				
	No	Yes	No	Yes	No	Yes	No	Yes	Value	df	Asymp. Sig.
Mind Map	228 89.4%	27 10.6%	83 77.6%	24 22.4%	21 70.0%	9 30.0%	38 97.4%	1 2.6%	19.212	3	**.000**
Free text	17 6.7%	238 **93.3%**	2 1.9%	105 **98.1%**	2 6.7%	28 **93.3%**	0 .0%	39 **100.0%**	6.059	3	.109
Word strings	115 45.1%	140 54.9%	30 28.0%	77 72.0%	5 16.7%	25 **83.3%**	15 38.5%	24 61.5%	15.700	3	**.001**
Full phrase	139 54.5%	116 45.5%	47 43.9%	60 **56.1%**	18 60.0%	12 40.0%	31 79.5%	8 20.5%	15.012	3	**.002**
Adv. Boolean operators	177 69.4%	78 30.6%	60 56.1%	47 43.9%	12 40.0%	18 **60.0%**	26 66.7%	13 33.3%	13.739	3	**.003**
Word stems	223 87.5%	32 12.5%	84 78.5%	23 21.5%	20 66.7%	10 **33.3%**	34 87.2%	5 12.8%	11.505	3	**.009**
Term truncation	221 86.7%	34 13.3%	87 81.3%	20 18.7%	21 70.0%	9 **30.0%**	34 87.2%	5 12.8%	6.654	3	.084
Wildcards	230 90.2%	25 9.8%	90 84.1%	17 15.9%	18 60.0%	12 **40.0%**	34 87.2%	5 12.8%	21.298	3	**.000**

Table 4.2.11.2 shows Pearson Chi-Square statistics for training on Web search and using Web search guides with each searching strategies.

Table 4.2.11.2: Pearson Chi-Square test results for Web searching strategies and Training on Web information search, Web search guides

Searching strategies	Pearson Chi-Square test results					
	Training on Web search			Web search guides		
	Value	df	Asymp. Sig.	Value	df	Asymp. Sig.
Mind Map	10.358	1	**.001**	11.576	1	**.001**
Free text	.455	1	.500	1.775	1	.183
Word strings	.273	1	.602	5.509	1	**.019**
Full phrase	2.548	1	.110	9.636	1	**.002**
Adv. Boolean operators	13.027	1	**.000**	5.965	1	**.015**
Word stems	7.159	1	**.007**	11.424	1	**.001**
Term truncation	26.192	1	**.000**	12.213	1	**.000**
Wildcards	8.163	1	**.004**	3.715	1	.054

Interpretations of the results including these two tables are organized separately for each strategy.

4.2.11.1 Mind maps or subject maps

According to table 4.2.11.1, a very lower percentage of participants use mind maps or subject maps before applying their search query and the two-sided asymptotic significance of the chi-square statistic (0.000) is less than 0.05, so it's safe to say that

there is a significant difference among the user category for usage of mind maps or subject maps. Further, it shows the usage of mind maps or subject maps is comparatively high for academic staff members (30%) and postgraduate students (22.4%). Further, according to table 4.2.11.2, there exists a significance difference among participation on training, using Web search guides on Web information search and using mind maps or subject maps before applying their search query.

Table 4.2.11.3: Mind map or subject map usage with training on Web search, Web search guides

		Training on Web search		Web search guides		Total
		No	Yes	No	Yes	
Mind map or Subject map	No	343	26	348	21	369
		87.5%	**68.4%**	87.4%	**65.6%**	85.8%
	Yes	49	12	50	11	61
		12.5%	31.6%	12.6%	34.4%	14.2%
Total		392	38	398	32	430
		100.0%	100.0%	100.0%	100.0%	100.0%

According to table 4.2.11.3, a very few respondents (14.2%) have applied mind map or subject maps before formulating their queries. Also, there is no any evidence to say that if someone has used Web search guides or has obtained any Web training their usage of mind map or subject map is higher. Although, 68.4% of respondents have obtained training on Web searching and 65.6% has used Web search guides, they have not used any mind maps or subject maps before formulating their search queries. It reveals that having training on Web searching or using Web search guides have not affected for the use of mind maps or subject maps.

4.2.11.2 Free text searching

According to table 4.2.11.1, most of the participants used free text searching method and there is no any evidence to indicate a difference among user category for fee text search.

Further, according to table 4.2.11.2, there is no evidence to show a significant difference in participation on formal training or using Web search guide in using free text searching.

4.2.11.3 Word strings

According to table 5.2.11.1, 83.3% of academic staff**,** 72% of postgraduate students, 61.3% of NDT students and 54.9% of undergraduate students used word strings for Web searching. Further, the two-sided asymptotic significance of the chi-square statistic (0.019) is less than 0.05; so it's safe to say that there is a significant difference among the user category for usage of word strings. In addition, according to table 4.2.11.2, there is a significant difference among using Web search guides on Web information search and using word string in their search query.

Table 4.2.11.4: Word strings usage with Web search guides

		Word strings		Total
		No	Yes	
Web search guides	No	158 39.7%	240 60.3%	398 100.0%
	Yes	6 18.8%	26 **81.2%**	32 100.0%
Total		164 38.1%	266 61.9%	430 100.0%

According to the table 4.2.11.4, 81.2% of respondents who has used Web search guides use word strings in their search queries. It reveals that if someone has used Web search guides their usage of word strings is higher.

4.2.11.4 Full phrase searching

According to table 4.2.11.1, 56.1% of postgraduate students, 45.5% of undergraduate students, 40% of academic staff members and 20.5% of NDT students have used full phrase in their search query. In addition, the two-sided asymptotic significance of the chi-square statistic (0.002) is less than 0.05, so it's safe to say that there is a significant

difference among the user categories for usage of full phrases. Further, according to table 4.2.11.2, there is a significant difference in using Web search guides on Web information search and using full phrases in their search queries.

Table 4.2.11.5: Full phrase search with Web search guides

		Full phrase		Total
		No	Yes	
Web search guides	No	225	173	398
		56.5%	43.5%	100.0%
	Yes	9	23	32
		28.1%	**71.9%**	100.0%
Total		234	196	430
		54.4%	45.6%	100.0%

According to the table 4.2.11.5, 71.9% of respondents who have used Web search guides, used full phrase search in their search queries. It reveals that if someone has used Web search guides their usage of word strings is higher.

4.2.11.5 Advanced Boolean operators

According to table 4.1.4.11.1, 60% of academic staff members, 43.9% of postgraduate students, 33.3% of NDT students and 30.6% of undergraduate students have used full phrase in their search query. In addition, the two-sided asymptotic significance of the chi-square statistic (0.003) is less than 0.05, so it's safe to say that there is a significance difference among the user category for usage full phrase. Further, according to table 4.2.11.2, there is a significant difference in participation on any training, using Web search guides on Web information search and using advanced Boolean operators in their search query.

Table 4.2.11.6: Usage of advanced Boolean operators with training on Web search, Web search guides

		Training on Web search		Web search guides		Total
		No	Yes	No	Yes	
Advanced Boolean operators	No	260 66.3%	14 36.8%	260 65.3%	14 43.8%	274 63.7%
	Yes	132 33.7%	24 **63.2%**	138 34.7%	18 **56.2%**	156 36.3%
Total		392 100.0%	38 100.0%		398 100.0%	32 100.0%

According to table 4.2.11.6, 63.2% of respondents have obtained any Web training and 56.2% of has used Web search guides. Also both categories have used Boolean operators in their search queries. It reveals that training on Web search or Web search guides have affected the usage of advanced Boolean operators in their search queries.

4.2.11.6 Word stems in Web searching

According to table 4.2.11.1, a very lower percentage of participants use word stems in their search query and the two-sided asymptotic significance of the chi-square statistic (0.009) is less than 0.05, so it's safe to say that there is a significant difference among the user category for usage of word stems. Further, it shows the usage of word stems is comparatively high for academic staff members (33.3%) and postgraduate students (21.5%). Further, according to table 4.2.11.2 there is a significant difference in participation on any training, using Web search guides on Web information searching and using word stems in their search query.

Table 4.2.11.7: Word stems usage with training on Web search, Web search guides

		Training on Web search		Web search guides		Total
		No	Yes	No	Yes	
Word stems	No	348 88.78%	12 31.60%	348 87.44%	12 37.50%	360 83.70%
	Yes	44 11.22%	26 **68.40%**	50 12.56%	20 **62.50%**	70 16.30%
Total		392 100.0%	38 100.0%	398 100.0%	32 100.0%	430 100.0%

According to table 4.2.11.7, 68.4% of respondents have obtained any training on Web searching and 62.5% of have used Web search guides and also the have used word stems in their search queries. It reveals that having training on Web searching or using Web search guides are affected for the usage of word stems in their search queries.

4.2.11.7 Term truncation in Web searching

According to table 4.2.11.1, a very lower percentage of participants have used term truncation in their search queries and there is no any evidence to show a difference among user categories for using term truncation. But, according to table 4.2.11.2 there is a significant difference in participation in some training, using Web search guides on Web information searching and using term truncation in their search query.

Table 4.2.11.8: Term truncation usage with training on Web search, Web search guides

		Training on Web search		Web search guides		Total
		No	Yes	No	Yes	
Term truncation	No	345 88.01%	17 44.70%	350 87.94%	12 37.50%	362 84.20%
	Yes	47 11.99%	21 **55.30%**	48 12.06%	20 **62.50%**	68 15.80%
Total		392 100.0%	38 100.00%	398 100.00%	32 100.00%	430 100.00%

According to table 4.2.11.8, 55.3% of respondents have obtained any training on Web searching and 62.5% of have used Web search guides and also the have used term truncation in their search queries. It reveals that having training on Web searching or using Web search guides affect the usage of term truncation in their search queries.

4.2.11.8 Wildcards in Web searching

According to table 4.2.11.1, a very lower percentage of participants use wildcards in their search query and the two-sided asymptotic significance of the chi-square statistic (0.000) is less than 0.05, so it's safe to say that there is a significant difference among the user categories in usage of wildcards. Further, it shows that the usage of wildcards is comparatively high for academic staff members (40%). Further, according to table 4.2.11.2 there is a significant difference between using Web search guides on Web information searching and using wildcards in their search query.

Table 4.2.11.9: Wildcards usage with training on Web search

		Training on Web search		Total
		No	Yes	
Wildcards	No	360 91.84%	11 28.90%	371 86.30%
	Yes	32 8.16%	27 **71.10%**	59 13.70%
Total		392 100.0%	392 100.00%	38 100.00%

According to table 4.2.11.9, 71.1% of respondents have obtained some training on Web searching and they have used wildcards in their search query. It reveals that having training on Web searching is affected for the usage wildcards in their search queries.

4.3 Query formulation

Respondents were given five search tasks and asked to search them on the Web and to provide information under topic familiarity, search statement he/she tried, time spent on each search task (maximum three attempts can be tried each task) and, whether they used advance features of the search engine and how they have judged the relevancy of information. The results were based on the queries formulated by 411 respondents for 5 search tasks (2055 queries, in total).

Four hundred eleven respondents (411) had formulated the initial query and out of them 88 respondents had reformulated query for the second time, but none of the respondents had tried a third time to reformulate their queries.

Table 4.3.1 shows the number of respondents and percentages who formulated initial and second queries for each user categories and their experience in using the search engines. It shows that about 20% undergraduates, 19% postgraduate students, 31% NDT students and 30% of academic staff members had reformulated their initial query. Overall, 21% of respondents have reformulated their initial query. Further table shows that the percentage of reformulating the query for the second time has decreased when increased the experience of using search engines. This shows that with increasing the experience of using search engines, percentage of reformulate initial query has been decreased.

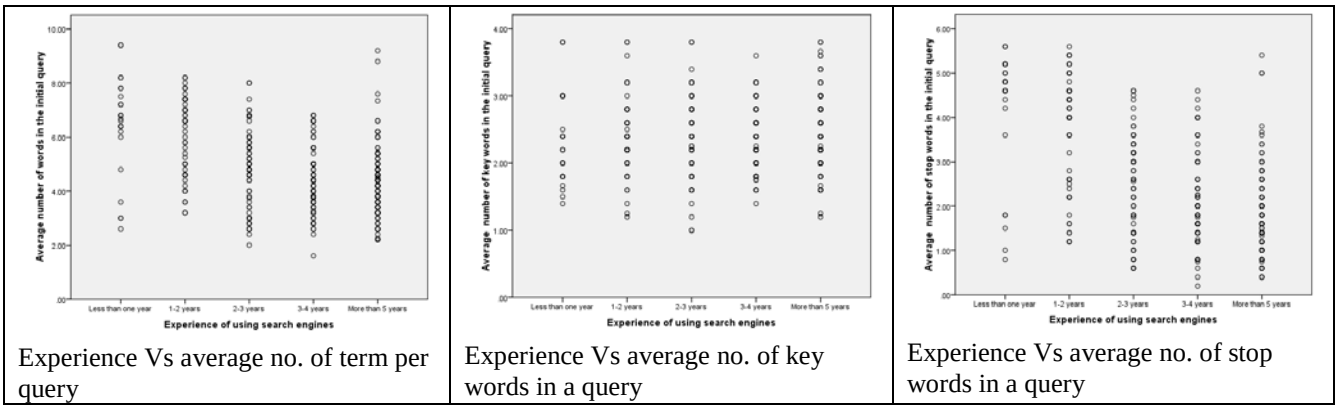
Table 4.3.1: Experience of using Web search engine for each user categories

Category		Experience of using search engines (years)					Total
		1>	1-2	2-3	3-4	5<	
Undergraduate Students	1st	19	43	56	75	46	239
		7.90%	18.00%	23.40%	31.40%	19.20%	100.00%
	2nd	7	9	17	10	5	48
		14.58%	18.75%	35.42%	20.83%	10.42%	100.00%
	Proportion	36.84%	20.93%	30.36%	13.33%	10.87%	20.08%
Postgraduate Students	1st	0	1	7	13	81	102
		0.00%	1.00%	6.90%	12.70%	79.40%	100.00%
	2nd	0	1	5	4	9	19
		0.00%	5.26%	26.32%	21.05%	47.37%	100.00%
	Proportion	0.00%	100%	71%	31%	11%	19%
Academic Staff Members	1st	2	0	1	0	27	30
		6.70%	0.00%	3.30%	0.00%	90.00%	100.00%
	2nd	2	0	1	0	6	9
		22.22%	0.00%	11.11%	0.00%	66.67%	100.00%
	Proportion	100%	0.00%	100%	0.00%	22.22%	30%
ITUM Students	1st	5	13	10	6	5	39
		12.80%	33.30%	25.60%	15.40%	12.80%	100.00%
	2nd	1	6	1	1	3	12
		8.30%	50.00%	8.30%	8.30%	25.00%	100.00%
	Proportion	20.00%	46.15%	10.00%	16.67%	60.00%	30.77%
Total	1st	26	57	74	94	159	410
		6.40%	13.90%	18.00%	22.90%	38.80%	100.00%
	2nd	10	16	24	15	23	88
		4.50%	14.80%	25.00%	20.50%	35.20%	100.00%
	Proportion	38.46%	28.07%	32.43%	15.96%	14.47%	21.46%

4.3.1 Initial query

In this part, length of the initial query, time spent for the formulation of the initial query and topic familiarities were analyzed.

4.3.1.1 Query length



Experience Vs average no. of term per query

Experience Vs average no. of key words in a query

Experience Vs average no. of stop words in a query

Figure 4.3.1.1 Relationship between the average query length, key words and stop words with search engine using the experience of respondents.

Figure 4.3.1.1 presents the search engine usage experience versus the average number of terms per initial query, average number of key words in an initial query and average number of stop word in initial query. It shows that with increasing experience of using search engines, the number of terms in an initial query decreased and the number of stops words in a initial query also decreased.

Table 4.3.1.1: Mean and stranded deviation of the initial query

Search engine experience (Years)	Average no. of terms in a initial query		Average no. of keywords in a initial query		Average no. of stop words in a initial query	
	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
1>	6.45	1.86	2.34	0.67	4.11	1.50
1-2	6.01	1.42	2.42	0.55	3.68	1.31
2-3	4.77	1.44	2.40	0.59	2.37	1.11
3-4	4.34	1.16	2.41	0.44	1.93	0.97
5<	4.17	1.14	2.47	0.48	1.70	0.86

Table 4.3.1.1 describes the mean and stranded deviation of average number of terms in an initial query, average number of keywords in an initial query and average number of stop

words in an initial query. It shows that respondents who were having less than one year experience in using search engines and their mean of average number of terms in an initial query has become higher (6.45) than others. This clearly shows the number of terms in an initial query was decreasing when experience of using search engines increased. Also respondents with less than one year experience of using search engines shows that the mean of average number of stop words in an initial query was higher (4.11) than others. Further it clearly shows the number of stop words in an initial query was decreasing when the experience of using search engines increased. But, the mean value of keywords in a query was nearly equal for each experience category.

Table 4.3.1.2: Pearson correlation experience of using Web search engine and initial query

		Average number of stop words in the initial query	Average number of words in the initial query	Average number of key words in the initial query
Experience of using search engines	Pearson Correlation	-.575	-.481	.063
	Sig. (2-tailed)	.000	.000	.200
	N	410	410	410

Table 4.3.1.2 describes the Pearson correlation coefficients for the experience of using search engines with average number of stop words, average number of words and average number of key words in an initial query. It clearly shows the experience of using search engine and average number of stop words in an initial query is negatively correlated (-0.575, 0.000) and also the average number of words in an initial query was negatively correlated (-0.481, 0.000). This means when a user becomes much experienced of using search engines, his/her usage of stop words has decreased. But there was no significant difference among search engine experience and using keywords. This may be because written queries describe most of the non experience users tended to type whole search statement. For example in search task 2 (Who is the vice president of the Nokia company in Finland), most of non expert users have typed “Who is the vice president of the Nokia

company in Finland?" But expert users have typed "vice president Nokia Company Finland". These both users have mentioned all keywords in their query, but non expert users' query length was higher because non expert users have used more stop words (ex. of, is, the, etc.).

4.3.1.2 Average time spent per query

Figure 4.3.1.2.1 presents the details of the time spent per query.

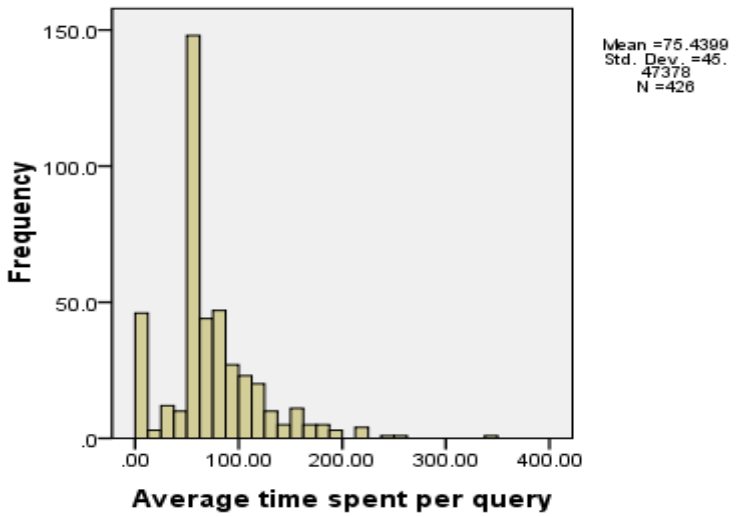


Figure 4.3.1.2.1: Average time spent per query

Figure 4.3.1.2.1 shows that average time spent per query was positively skewed and its mean value was 75.4399 seconds.

Figure 4.3.1.2.2 summarizes the relationship between average time spent per query and web experience. It reveals that the time spent to formulate a query has increased when web experience increased. That means the experienced users have taken more time and create queries. Pearson's correlation coefficient further confirmed this relationship by testing the hypothesis;

H_0: There is no relationship between web experience and time spent per query

Versus

H_1: There is a relationship between web experience and time spent per query

Results are shown in table 4.3.1.2.1.

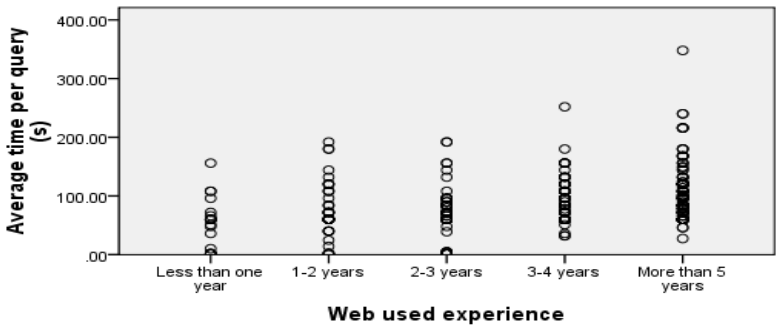


Figure 4.3.1.2.2: Scatter plot of average time spent per query vs web experience

Table 4.3.1.2.1: Pearson correlation between average time spent per query and web experience

		Web used experience
Average time per query	Pearson Correlation	.330**
	Sig. (2-tailed)	**.000**
	N	425

**. Correlation is significant at the 0.01 level (2-tailed).

According to table 4.3.1.2.1 the significant value of the test was 0.000 and Pearson correlation value was 0.214. It indicates that the null hypothesis can be rejected at 0.05 significant level. Thus, when the web experience has increased, time spent per query was also increasing.

4.3.1.3 Topic familiarity

Respondents were asked to state the topic familiarity for each search task before formulating a query. The scale used for the topic familiarity was; 1 – I don't know the topic at all; 2 – I know a bit about the topic; 3 – I know somewhat about the topic; 4 – I know the topic; 5 – I know the topic very well. Table 4.3.1.3.1 summarizes the mean time spent per query for each familiarity level and ANOVA results for each search task.

Table 4.3.1.3.1: Mean time spent per query for each familiarity level and ANOVA results

Time spent (in second)							
Search Task	**Mean time (Topic familiarity)**					**ANOVA Results**	
	1	**2**	**3**	**4**	**5**	**F**	**Sig.**
ST1	122.63	91.37	73.31	62.70	46.75	27.843	.000
ST2	165.43	119.69	80.71	64.39	51.11	94.339	.000
ST3	106.00	96.02	92.38	73.57	81.62	3.495	.008
ST4	136.91	101.64	74.36	66.07	67.83	35.508	.000
ST5	112.04	93.43	95.22	85.47	68.57	2.812	.025

It reveals that the mean time spent per query differed according to the topic familiarity. This was significant for all search tasks under 0.05 significance level. Further, it shows when a topic was familiar to a user; the spending time to formulate a query was decreasing.

Table 4.3.3.2 presents the Pearson correlation coefficients (r) for all search tasks.

Table 4.3.1.3.2: Pearson correlation between times spent per query vs topic familiarity

Pearson Correlation			
Search Task	**Topic familiarity vs time spent per search query**		
	r	**Sig. (2-tailed)**	**N**
ST1	-.456	0.000	402
ST2	-.680	0.000	408
ST3	-.150	0.002	404
ST4	-.468	0.000	411
ST5	-.155	0.002	402

It revealed a negative correlation between topic familiarity and time spent to formulate a query under 0.05 significance level. It means when search topic was much familiar, the time spent per query has been decreased.

Table 4.3.3.3 shows the number of participants who have used synonyms and their topic familiarity for each search task. Also table 4.3.3.4 shows that Pearson Chi-Square results for topic familiarity vs synonyms used. These revealed a significant relationship between topic familiarly and synonyms used.

Table 4.3.3.3: Number of users who used synonyms and topic familiarity for each search task

Search Task	Topic familiarity					Total
	1	2	3	4	5	
ST1	0 .0%	3 21.4%	3 21.4%	6 42.9%	2 14.3%	14 100.0%
ST2	0 .0%	1 14.3%	2 28.6%	3 42.9%	1 14.3%	7 100.0%
ST3	0 .0%	7 6.9%	8 7.8%	67 65.7%	20 19.6%	102 100.0%
ST4	4 10.8%	8 21.6%	5 13.5%	15 40.5%	5 13.5%	37 100.0%
ST5	0 .0%	0 .0%	0 .0%	2 40.0%	3 60.0%	5 100.0%

Table 4.3.3.4: Pearson Chi-Square test results for topic familiarity vs synonyms used

Search task	Value	df	Asymp. Sig. (2-sided)
ST1	13.971	4	.007
ST2	15.492	4	.004
ST3	49.812	4	.000
ST4	16.353	4	.003
ST5	21.941	4	.000

4.3.2 Secondly formulated query

Table 4.3.2.1: Pearson correlation coefficient for second query

		Average length of second query	Average number of key words in second query	No of stop words in second query	Average spend time for second query
Experience of using search engines	Pearson Correlation	-.097	.086	-.160	-.044
	Sig. (2-tailed)	.367	.427	.137	.681
	N	88	88	88	88

Table 4.3.2.1 describes the Pearson correlation coefficient for experience of using search engines with average length of second query, average number of keywords in the second query, average number of stop words in the second query and average time spent on second query. According to significant values, there was no evidence to say that there was

a positive or negative correlation with experience of using search engines and the average length of the second query, average number of keywords in the second query, average number of stop words in the second query or average time spent to formulate the second query.

4.3.4 Search interface

Table 4.3.4.1 presents the web searching interface used by each user category.

Table 4.3.4.1: Web interface use each user category

		Web search Interface			Total
		Simple search	Advance search	Simple & Advance search	
Category	Undergraduate Students	208 83.5%	33 13.3%	8 3.2%	249 100.0%
	Postgraduate Students	77 72.6%	17 16.0%	12 11.3%	106 100.0%
	Academic Staff Members	22 73.3%	1 3.3%	7 23.3%	30 100.0%
	NDT Students	28 75.7%	9 24.3%	0 .0%	37 100.0%
Total		335 79.4%	60 14.2%	27 6.4%	422 100.0%

Table shows that most of the respondents have used a simple search interface. The relevant percentage was 79.4%. It further revealed that undergraduate and NDT students were the groups who used simple search interface mostly. The relevant percentages were 83.5% and 75.79%.

Table 4.3.4.2 depicts the Chi Square test results between the user categories and web search interface used by them.

Table 4.3.4.2: Chi-Square test results between web interface and user category

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	30.696	6	**.000**
Likelihood Ratio	28.372	6	.000
Linear-by-Linear Association	4.368	1	.037
N of Valid Cases	422		

It shows that there was a significance relationship between web interfaces used as user categories.

Table 4.3.4.3 summarizes the criteria of judging relevancy of information by web experience.

Table 4.3.4.3: Criteria of judging relevancy of information by web experience

Criteria of judge	**Web experience (years)**					**Total**
	<1	**1-2**	**2-3**	**3-4**	**>5**	
By Title	9 (37.50%)	28 (45.16%)	29 (36.71%)	53 (57.61%)	82 (47.13%)	201 **(46.64%)**
Highlighted words	4 (16.67%)	18 (29.03%)	28 (35.44%)	43 (46.74%)	58 (33.33%)	151 (35.03%)
Reading descriptions	8 (33.33%)	17 (27.42%)	31 (39.24%)	34 (36.96%)	86 **(49.43%)**	176 (40.84%)
Reading URLs	1 (4.17%)	2 (3.23%)	5 (6.33%)	8 (8.70%)	15 (8.62%)	31 (7.19%)

Table 4.3.4.3 reveals that most of the respondents have judged relevancy by checking title (46.64%) of their search results and reading descriptions (40.84%). But substantial respondents (35.03%) considered highlighted words. Further, it reveals that the percentage of the respondents who read the descriptions has increased with their long

experience of using search engines.. It means that a user who has become an expert in web has read more descriptions of their search results.

4.4 Observation Data

A total of 40 users were observed among the four main user categories of University of Moratuwa. Table 4.4.1 depicts the response rate of each user category separately.

Table 4.4.1: Response rate of different user categories

	Faculty/Institute				
Category	Arch	Eng	IT	NDT	Total
Undergraduate Students	3 30.0%	3 30.0%	4 40.0%	0 .0%	10 100.0%
Postgraduate Students	2 20.0%	6 60.0%	2 20.0%	0 .0%	10 100.0%
Academic Staff Members	3 30.0%	2 20.0%	3 30.0%	2 20.0%	10 100.0%
NDT Students	0 .0%	0 .0%	0 .0%	10 100.0%	10 100.0%
Total	8 20.0%	11 27.5%	9 22.5%	12 30.0%	40 100.0%

According to table 4.4.1, 40% of IT undergraduate students, 60% of engineering postgraduate students, 30% of Architecture academic staff members and 30% of IT academic staff members were subjected for the observation.

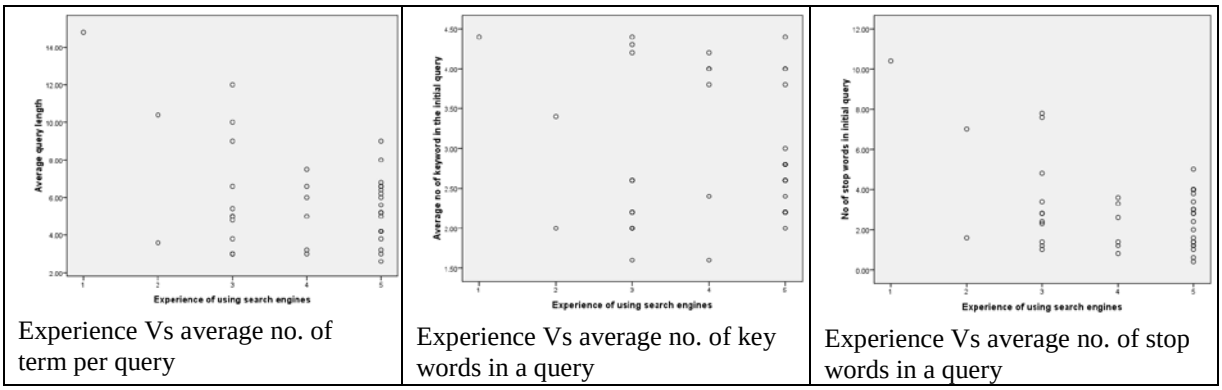
Table 4.4.2: Experience of using Web search engine for each user categories

Category		Experience of using Web search engines					
		1>	1-2 years	2-3 years	3-4 years	5 <	Total
Undergraduate Students	1st Query	0	0	7	1	2	10
		.0%	.0%	70.0%	10.0%	20.0%	100.0%
	2nd Query	0	0	7	0	0	7
		.0%	.0%	100.0%	.0%	.0%	100.0%
	Proportion	.0%	.0%	100.0%	.0%	.0%	70.0%
Postgraduate Students	1st Query	0	0	2	2	6	10
		.0%	.0%	20.0%	20.0%	60.0%	100.0%
	2nd Query	0	0	1	0	2	3
		.0%	.0%	33.3%	.0%	66.7%	100.0%
	Proportion	.0%	.0%	50.0%	.0%	33.33%	30%
Academic Staff Members	1st Query	0	0	1	0	9	10
		.0%	.0%	10.0%	.0%	90.0%	100.0%
	2nd Query	0	0	1	0	0	1
		.0%	.0%	100.0%	.0%	.0%	100.0%
	Proportion	.0%	.0%	100.0%	.0%	.0%	10%
NDT Students	1st Query	1	2	1	3	3	10
		10.0%	20.0%	10.0%	30.0%	30.0%	100.0%
	2nd Query	1	2	1	2	1	7
		14.3%	28.6%	14.3%	28.6%	14.3%	100.0%
	Proportion	100.0%	100.0%	100.0%	66.66%	33.33%	70%
Total	1st Query	1	2	11	6	20	40
		2.5%	5.0%	27.5%	15.0%	50.0%	100.0%
	2nd Query	1	2	10	2	3	18
		5.55%	11.11%	55.55%	11.11%	16.66%	100.0%
	Proportion	100%	100%	90.91%	33.33%	15%	45%

Table 4.4.2 describes the number of respondents who formulated both first and second queries and the experience of using search engines for all user categories. It shows 20% of undergraduate, 30% of NDT students, 60% of postgraduate students and 90% of academic staff members were having more than 5 years experience using search engines. Also 70% of undergraduate and NDT students, 30% of postgraduate students and 10% of academic staff members have reformulated their queries. However, 45% of respondents

have reformulated their queries. About 90% of non experience users (less than 3 years experience using web search engine) have also reformulated queries.

4.4.1 Initial query



Experience Vs average no. of term per query

Experience Vs average no. of key words in a query

Experience Vs average no. of stop words in a query

Figure 4.4.1.1 Relationship between the average query length, key words and stop words with search engine using experience of respondents.

Figure 4.4.1.1 presents the search engine using experience versus average number of terms per initial query, average number of key words in an initial query and average number of stop word in an initial query. It shows that when the experience of using search engines increased, the number of terms in an initial query has been decreased and the number of stops words in an initial query has also been decreased.

Table 4.4.1.2: Pearson correlation experience of using Web search engine and initial query

		Average query length in initial query	Average no of keyword in the initial query	No of stop words in initial query	Average precision for initial query	Average time spend for initial query
Experience of using search engines	Pearson Correlation	-.399	-.074	-.465	.852	.383
	Sig. (2-tailed)	.011	.648	.002	.000	.015
	N	40	40	40	40	40

Table 4.4.1.2 describes that Pearson correlation coefficients for the experience of using search engines with average number of stop words, average number of words, average

number of key words in an initial query and average precision for initial query. It clearly shows that the experience of using search engines and average number of stop words in the initial query were negatively correlated (-0.465, 0.002), in addition, the average number of words in the initial query were negatively correlated (-0.399, 0.011). When the user has become more experienced in using search engines, usage of stop words have been decreased. But there was no significant difference among search engine experience and using keywords. Also the average precision of initial query and the experience of using search engines were highly and positively correlated (0.852, 0.000). With increasing of search engine usage experience, his/her precision of initial query also increased. Further, with increasing the experience of using search engines, time spent to create a query has also been increased (0.383, 0.015).

4.4.2 Secondly formulated (reformulated) query

Table 4.4.2.1: Pearson correlation coefficient for second query

		Avequerylength2	Average no of keyword in second query	No of stop words in second query	Average spend time for second query	Average precision for second query
Experience of using search engines	Pearson Correlation	.428	.310	.320	-.051	**.608**
	Sig. (2-tailed)	.077	.211	.195	.841	**.007**
	N	18	18	18	18	18

Table 4.4.2.1 describes Pearson correlation coefficients for the experience of using search engine with average length of second query, average number of keywords in the second query, average number of stop words in the second query and average time spent for second query. According to significant values, there was no evidence to say that there was a positive or negative correlation with the experience of using search engines and average

length of second query, average number of keywords in the second query, average number of stop words in the second query or average time spent to formulate the second query. But there was a positive correlation with the experience of using search engines and the average precision of the second query (0.608, 0.007).

4.5 Validation of the survey results

This section presents the comparison of observation and survey data. Levene's test was used for Equality of Variances and t-test was used for equality of mean. These two tests were used to validate survey data compared with observation data.

Table 4.5.1: Mean and standard deviation of average number of words in the initial query for each user category.

	Category	Type	N	Mean	Mean Diff	Std. Deviation	Std. Dev. Diff
Average number of words in the initial query	UG	Survey	24	4.4730	-0.107	1.28318	0.56238
		Observation	10	4.5800		.72080	
	PG	Survey	10	3.9647	-0.5153	1.08716	-.6364
		Observation	10	4.4800		1.72356	
	Staff	Survey	30	4.4800	-0.76	1.32935	-.49103
		Observation	10	5.2400		1.82038	
	NDT	Survey	39	4.2530	-0.28	1.75813	0.20007
		Observation	10	4.5330		1.55806	
	All	Survey	41	4.3265	-0.3817	1.30760	-.18515
		Observation	40	4.7082		1.49275	

Table 4.5.1 describes mean and stranded deviation of the average number of words in the initial query for survey and observation data for all user categories. It shows that the

difference of standard deviation and mean value of survey and observation method was very low (less than 0.8) for all user categories.

Table 4.5.2: Levene's and t-test values for average initial query length for the observed and survey data.

			Levene's Test for Equality of Variances		t-test for Equality of Means						
			F	Sig	t	df	Sig	Mean Difference	Std.	95% Confidence Interval of the Difference	
										Lower	Upper
Average number of words in the initial query	UG	EV	2.700	**.262**	-.262	248	**.79**	-.107	.408	-.912	.698
		NEV			-.441	11.53	.66	-.107	.242	-.637	.423
	PG	EV	10.78	.001	-1.35	110	.18	-.515	.381	-1.272	.241
		NEV			-.928	9.714	**.37**	-.515	.555	-1.758	.727
	Staff	EV	1.609	**.212**	-1.42	38	**.16**	-.76	.533	-1.839	.319
		NEV			-1.22	12.36	.24	-.76	.624	-2.116	.596
	NDT	EV	.167	**.684**	-.459	47	**.64**	-.28	.610	-1.507	.947
		NEV			-493	15.44	.62	-.28	.567	-1.486	.926
	All	EV	1.563	**.212**	-1.74	449	**.08**	-.381	.219	-.812	.049
		NEV			-1.56	45.01	.12	-.381	.244	-.874	.111

EV- Equal variance assumed, NEV - Equal variance not assumed

Table 4.5.2 describes Levene's and t-test values for observation and survey data for length of the initial query. It shows there was no significant difference (under 0.05 significant level) between observation and survey data in the variance for staff, undergraduate and NDT students. But there was a significant difference between observation and survey data in the variance for postgraduate students. However, assuming or not assuming equal variance t-test results shows that there was no significance mean difference in the initial query length (under 0.05 significant level) between observation and survey data.

Table 4.5.3: Levene's and t-test values for average second query length for the observed and survey data.

			Levene's Test for Equality of Variances		t-test for Equality of Means						
			F	Sig	t	df	Sig	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
										Lower	Upper
Average number of words in the initial query	UG	EV	.763	**.386**	-.971	56	**.336**	-.5101	.525	-1.562	.5418
		NEV			-.822	11.36	.428	-.5101	.621	-1.871	.8507
	PG	EV	1.134	**.297**	-.441	24	**.663**	-.331	.749	-1.877	1.215
		NEV			-.390	8.769	.706	-.331	.848	-2.258	1.596
	Staff	EV	.149	**.706**	-.425	12	**.679**	-.444	1.046	-2.724	1.835
		NEV			-.425	8.415	.681	-.444	1.044	-2.833	1.944
	NDT	EV	.100	**.756**	.394	16	**.699**	.3977	1.009	-1.742	2.537
		NEV			.396	10.21	.700	.3977	1.005	-1.836	2.631
	All	EV	1.192	**.277**	-1.14	114	**.253**	-.4255	.370	-1.159	.3087
		NEV			-1.08	41.71	.284	-.4255	.392	-1.217	.3666

EV- Equal variance assumed, NEV - Equal variance not assumed

Table 4.5.3 presents Levene's and t-test values for length of the second query according to observation and survey data. It shows that there was no significance mean or variance difference (under 0.05 significant level) between observation and survey data in the length of second query for any user category.

According to the above results (table 4.5.2 and 4.5.3), It can be said that the data collection methods have not effected for the length of the query.

Table 4.5.4 presents Levene's test values and t-test values for time spent to formulate the initial query. According to significant value of the Levene's test, there was no variance difference in the time spent to formulate a query, between survey and observation data for all user categories. Also t-test values show that there was no mean difference in the time spent to formulate a query, between survey and observation data for all user categories.

The significant values clearly show that there was no significant difference between survey and observation data for the time spent to formulate an initial query.

Table 4.5.4: Levene's and t-test values for average time spend for the initial query for the observed and survey data.

			Levene's Test for Equality of Variances		t-test for Equality of Means						
			F	Sig	t	df	Sig	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
										Lower	Upper
Average number of words in the initial query	UG	EV	2.154	**.143**	-.479	248	**.632**	-6.526	13.61	-33.33	20.281
		NEV			-.320	9.314	.756	-6.526	20.36	-52.36	39.313
	PG	EV	3.973	**.049**	-1.17	110	**.241**	-14.82	12.58	-39.76	10.11
		NEV			-.854	9.82	.413	-14.82	17.35	-53.60	23.95
	Staff	EV	.200	**.657**	.167	38	**.868**	3.300	19.75	-36.69	43.29
		NEV			.208	24.85	.837	3.300	15.82	-29.31	35.91
	NDT	EV	.171	**.681**	.026	48	**.979**	.3800	14.62	-29.02	29.78
		NEV			.031	17.67	.976	.3800	12.41	-25.73	26.49
	All	EV	1.434	**.232**	-.788	450	**.431**	-5.517	7.004	-19.28	8.248
		NEV			-.709	45.06	.482	-5.517	7.780	-21.18	10.152

EV- Equal variance assumed, NEV - Equal variance not assumed

Table 4.5.5 presents Levene's test values and t-test values for the time spent to reformulate the query. According to the significant values of the Levene's test, there was no variance difference in the time spent to reformulate a query, between survey and observation data for all user categories. Also t-test values show that there was no mean difference in the time spent to reformulate a query, between survey and observation data for all user categories. Values in the table further clearly show that there was no significant difference between survey and observation data for the time spent to reformulate the initial query.

Table 4.5.5: Levene's and t-test values for average time spend for reformulate the query for observed and survey data.

			Levene's Test for Equality of Variances		t-test for Equality of Means						
			F	Sig	t	df	Sig	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
										Lower	Upper
Average number of words in the initial query	UG	EV	1.407	**.241**	-.884	56	**.380**	-19.17	21.69	-62.633	24.279
		NEV			-.704	10.94	.496	-19.17	27.22	-79.129	40.775
	PG	EV	.419	**.524**	-.283	24	**.780**	-3.413	12.08	-28.346	21.519
		NEV			-.336	15.83	.741	-3.413	10.15	-24.961	18.134
	Staff	EV	.530	**.481**	-.094	12	**.927**	-1.733	18.43	-41.897	38.431
		NEV			-.090	7.33	.931	-1.733	19.32	-47.01	43.545
	NDT	EV	1.665	**.215**	.207	16	**.839**	2.500	12.09	-23.149	28.149
		NEV			.245	15.33	.809	2.500	10.18	-19.167	24.167
	All	EV	.050	**.823**	-.566	114	**.572**	-6.022	10.63	-27.096	15.051
		NEV			-.534	41.58	.597	-6.022	11.28	-28.81	16.765

EV- Equal variance assumed, NEV - Equal variance not assumed

Table 4.3.1.2 and 4.4.1.2 show that with the increased experience of using Web search engines, the usage of stop words in the initial query was decreasing and time spent to formulate an initial query was increasing for the both data collected methods, observation and survey. According to table 4.3.2.1 and 4.4.2.1 there was no relationship between query reformulation in both data collected methods. Therefore, these results imply that observation data shows the same characteristics of the survey data.

Chapter 5 – Conclusion and Recommendations

5.1 Conclusion

The focus of this study was to evaluate the factors affecting query formulation in Web information searching of the academic community in the University of Moratuwa. It was expected to identify the purposes for searching the web, investigate factors affecting on formulating queries, identify the types of Web information search by the academic community, and to examine how the academic community conceptualizes searching strategies.

There was a relationship between the Web experience & search engine experience as well as the Web experience & computer experience for undergraduates, NDT students and postgraduate students. The test results showed in the table 4.2.8.2 and 4.2.1 proves that the pattern was significant. Search engine experience has increased with the Web experience increased and Web experience has increased when computer experience increased.

5.1.1 Purpose for searching the Web

The ranking of student responses on the purpose of their Web surfing was validated using the Friedman test. Here, the lowest rank is for the highest priority, and vice versa.

According to the test, the purpose of surfing web is different among different user groups. For students (NDT, undergraduate and postgraduate), the purpose is mainly sending emails. But for staff, the main purpose of using Web is for research activities. Among students there is a significant difference between faculties as well. Architecture and

Engineering students' first preference is mostly to use the Web to send or receive e-mails. But IT and NDT students' mainly use it for education purposes.

There is a striking difference between the year/level of study too. Although Table 4.1.4.4.4 put forward the results of all years, we consider years 1-3 because all faculties of UOM and the ITUM mainly represent the samples of the first three years. Second year students are mostly using Web for education purposes and third year students are mostly using Web to send or receive e-mails. However, the fourth year students use Web for entertainment purposes.

There is a gender difference as well. Although, most of the male undergraduate students are using web to send or receive e-mails, female students' first preference is for 'education purposes'. At this level, we are not in a position to explain this behavior or grade the activities as good or bad. Therefore, it is open for future studies.

Postgraduate students are mostly homogenous and they mainly use the Web for sending and receiving e-mails. Finding information for research projects is only the second preference.

A similar pattern is evident among academic staff. Overall, their first priority is to find information for research activities. However, there is a difference among senior and junior faculty members. Senior lecturers primarily use Web to searching information for research purposes and their second priority is to send or receive e-mails. But probationary lectures' priorities are in the reverse direction.

5.1.2 Factors affecting query formulation

There is a relationship between number of stop words in the query and the web experience (Figure 4.1.5.2). The test in Table 4.1.5.1.2 proves that the pattern is significant. Number of stop words in the query is decreasing when Web experience of users increases. Empirical analysis reveals that higher the period of experience results in an increased practice of omitting stop words from the queries formulated. In personal discussions with respondents, they reveal that the normal practice of novices is to use the complete question as the query.

The time spent to formulate a query increases as the Web experience increases. This means that experienced user take more time to create queries.

When the search topic is much familiar to the user, the time spent per query decreases.

There isn't any significance in difference in using synonyms by respondents having different experience. However, it was observed that there is a strong link between using synonyms and the familiarity of the subject. Those who have a higher level of familiarity on the subject use synonyms in queries much more often than others.

Finally, It was revealed that there is no relationship between query reformulation in both data collected methods (table 4.3.2.1 and 4.4.2.1). Therefore, it is implied that observation data shows the same characteristics of the survey data.

5.1.3 Conceptualizing search strategy

Usage of mind maps or subject maps is comparatively higher among academic staff members (30%) and postgraduate students (22.45%). Only 10.6 % of undergraduates use mind mapping techniques in this stage of query formulation. Among NDT students, the number is even lower. Therefore, it is concluded that mind mapping is not an established activity during query formulation by the selected population.

A larger percentage of population, irrespectively of the group, has used free text searching. According to table 4.2.11.1, it is over 90%. Among those who do free text searching, a larger percentage hasn't received training in searching techniques.
However, it is noted that modern the popular search engines (Eg: Google) support free text searching more and more. For example, Google automatically insert AND operators and quotation marks for phrase searching where necessary.

Most of the academic staff members (83.3%) and postgraduate students (72%); and a considerable number of NDT (61.3%) and undergraduate students (54.9%) are using word strings method. Further, it is evident that those who use web search guides have a higher use of word strings (Table 4.2.11.1).

A considerable percentage of postgraduate students (56.1%), undergraduate students (45.5%), and academic staff members (40%) have used phrases in their search queries. But a very lower percentage of NDT students (20.5%) have used full phrases. There is no significant difference between those who have had a training on searching techniques and those who haven't. However, there is a difference between those who have used search

guides and those who haven't. Users who had used guides have used phrases in search queries than the rest of the population (Table 4.2.11.5).

Sixty per cent of academic staff members, 43.9% of postgraduate students, 33.3% of NDT students and 30.6% of undergraduate students have used full phrases in their search queries. Further, if someone has used web search guides or has obtained any formal training on Web searching, their usage of Boolean operators is higher (Table 4.2.11.1).

A very lower percentage of participants use word stems in their search query. But, the usage of word stems is comparatively high for academic staff members (33.3%) and postgraduate students (21.5%). Further, if someone has used web search guides or has obtained any formal training on Web searching, their usage of word stems is higher.

An even lower percentage of participants have used term truncation in their search query, and there is no any evidence to indicate a difference among user category for using term truncation. But, if someone has used web search guides or obtained any formal training on Web searching, their usage of term truncation is higher.

Usage of wildcards is comparatively high for academic staff members (40%) and of those who have used web search guides; their usage of wildcards for their search queries is higher.

There is a clear difference of skills on search techniques among different user categories. In general, the following pattern of skill levels can be seen in decreasing order.

Academic staff >> **Postgraduate Students** >> **Undergraduate/Diploma students**

5.1.4 Type of Web information search

In average, almost all respondents download e-books from web searches. The response rate is 67.4% (Table 4.2.5.1.2). The other most common resource they download is reference sources (Eg. Encyclopedias, dictionaries etc.). This is more than 60%. Use of periodicals is low among the total population.

Among different entities, highest use of e-books is from the Faculty of Information Technology (96.7%). The result is comparable with their IT skills. Using Computers and Internet is a routine activity of this student group.

Among NDT students this number is lower (Table 4.2.5.1.4). Instead, they use newspapers as the mostly used resource type. This can be explained as a result of lower demand of primary sources from this student group (diploma level).

Use of information sources also demonstrate some interesting findings. Almost all strata of the population have the highest affinity to use social networks than other sources

(Table 4.2.5.1.3). Between different entities, there are some differences in using social networks (Table 4.2.5.1.5). Here, the IT faculty students have a higher tendency to use social networks than all other entities (86.2%). The Faculty of Architecture, Engineering, and NDT students use social networks in decreasing order (75.8%, 68.0%, and 54.3% respectively). This cannot be explained using a single phenomenon. Therefore, the following reasons can be suggested for the differences.

Table 5.1.4.1: Summary of the usage of Web resources/sources

Entity	**Percentage**	**Observation**	**Reason**
Faculty of IT	86.2	Higher than all others	Use of computers and Internet is their routine activity.
Faculty of Architecture	75.8	High	Have access to computers and Internet
Faculty of Engineering	68.0	High, but lower than faculty of Architecture.	Have access, but have a comparatively higher workload (empirical observation). Therefore, unable to spend much time on Internet
NDT students	54.3	Low	Lesser opportunities to access computers and Internet, because only a fewer percentage have computers at home (Fig.4.1.3.1)

When considering postgraduate students from all three faculties, their preferences and selection is almost similar to undergraduates. It is observed that they also use social networks as the most preferred information source. However, the pattern of using information sources and resources among postgraduate students has a slight difference. As expected, they use periodicals much more than other sources (Table 4.2.5.2.2). There is a small deviation of this result among postgraduate students from Faculty of IT. Here, 100% of them use e-books while 93.3% use e-journals/magazines.

Faculty-wise distribution of use of information sources and resources have slight differences. Fifty one per cent of Architecture students use ‘databases’, while 65.1% of Engineering students use ‘social networks’. Eighty two point two per cent of IT students use Web portals.

Use of resources and sources of academic staff members is somewhat similar to postgraduate students. A 90% of academic staff members use journals/magazines and 89.3% use databases. Again, academic staff from the Faculty of IT is an offset. Hundred per cent of academic staff members from the Faculty of Information Technology are using Web portals, databases, social networks and subject gateways as information sources.

The overall result suggests that a large percentage of students use social networks as the mostly used information source, while the academic staff uses databases as the main source. This exemplifies the tendency of younger generation in using social networks as a source. Their affinity to proper/traditional sources is much lower.

5.2 Recommendations

University academic community is spending more and more time creating, seeking, retrieving and using electronic information. But their interactions with Web search engines are not much efficient and effective. They are not having maximum beneficial with the features of Web search engines. To adjust to these factors and to human behavior we need a new generation of Web searching tools that work with people to help them persist in electronic information seeking to resolve their information problems.

Study also reveals that students are not taking full benefit of the proper electronic sources provided by the library and resources/sources freely available in Web. Therefore it is necessary to conduct training programs on available electronic information sources and resources in the library and on Web as well.

In addition, study reveals that not much usage of the searching strategies and conceptualizing search query with mind mapping techniques are visible of the students. Therefore those strategies should be incorporated into their learning programs in GPA[1] or non-GPA courses. This could be highlighted in the proposed 'library module'[2].

Most of the academic staff members also do not gain the full benefit of the search engines using their advance operating techniques. Therefore library can organize training programs/workshops for academic staff members to introduce Web searching techniques.

In general, study reveals that most of the respondents except their categories possess a very poor knowledge in Web information searching. Therefore, library can develop web portals and virtual libraries for each specific subject to be easily accessed the electronic information.

Further the study reveals that most of the respondents compel to use the free text searching. Therefore it is recommended to use intelligence tools in designing search engines.

[1] Grade Point Average
[2] A non-GPA course proposed to the Faculty of Engineering.

Bibliography

Anick, P. (2003) Using terminological feedback for web search refinement: a log-based study. *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, p.88-95.

Aula, A. (2003), Query formulation in web information search. *Proceedings of the IADIS international conference WWW/Internet 2003*, Algarve, Portugal, November 2003. IADIS Press, Lisbon, Portugal, p. 403–410

Aula, Anne and Kaki Mika (2003) Understanding expert search strategies for designing user-friendly search interfaces. [Online], Available from: http://www.iadis.net/dl/final_uploads/200302C098.pdf [25th March 2010].

Aula, A. (2005) Studying user strategies and characteristics for developing web search interfaces. Academic dissertation. University of Tampere.

Barsky and Bar-Ilan (2005) From the search problem through query formulation to results on the web. *Online information review*, 29(1), p.75-89.

Bates, M.J. (1989) The design of browsing and berrypicking techniques for the on-line search interface. *Online Review*, 13**(**5), p.407–431.

Belkin, N.J. et al. (1995) Cases, scripts and information seeking strategies: on the design of interactive information retrieval systems. *Expert Systems with Applications*, 9, p.379–395.

Belkin, N. J. (2000) Helping people find what they don't know. *Communications of the ACM*, 43(8), p.58-61.

Bhatti, Rubina. (2010) Information needs and information- seeking behavior of faculty members at the Islamia University of Bahawalpur. Available from: http://www.webpages.uidaho.edu/~mbolin/bhatti.htm [Accessed 30th Dec 2010].

Borgman, Christine L. (1987) The study of user behaviour on information retrieval systems. *SIGCUE* [online], 19(3-4) Springer/Summer 1987. p.35-48. Available from: http://delivery.acm.org/10.1145/970000/964086/p35-borgman.pdf?key1=964086&key2=8272247921&coll=DL&dl=ACM&CFID=9591540&CFTOKEN=53892468 [Accessed 4th March 2010].

Ciolek. T.M. (2003) Asian Studies Online-a timeline of major developments [Online], Available: http://coombs.anu.edu.au/asian-studies-timeline.html [27th Dec. 2009].

Cleveland, D. B and Cleveland, A.D. (eds) (1990) Introduction to Indexing and Abstracting. Englewood, CO: Libraries Unlimited, Inc.

Daniel, Wayne W. (2000) *Bio statistics: a foundation for analysis in the health science*. John Wiley & Sons Inc. p. 130.

De Silva, K.M. and Peiris, G.H. (1995) The University system of Sri Lanka. International Center for Ethnic Studies , Kandy.

Fenichel, C. H. (1981) Online searching: Measures that discriminate among users with different types of experiences. *Journal of the American Society for Information Science*, 32(1), p.23-32.

Fidel, R. and Soergel, D. (1983) Factors affecting online bibliographic retrieval: A conceptual framework for research. *Journal of the American Society for Information Science*,34(3), p.163-180.

Fields, B., Keith, S., & Blandford, A. (2004) Designing for expert information finding strategies. *Proceedings of HCI*, p. 89-102.

Glenn, D. Israel. Determining sample size. University of Florida IFAS Extension. Available from: http://edis.ifas.ufl.edu. [Accessed 2nd Nov 2010]

Hölscher, C., and Strube, G. (2000) Web search behavior of internet experts and newbies. *Proceedings of the 9th International WWW conference*, p. 337-346.

Hsieh-Yee, Ingrid. (2001) Research on Web search behavior. *Library and Information Science Research*, 23, p.167-185.

Hsieh-Yee, Ingrid. Effects of Search Experience and Subject Knowledge on Online Search Behavior: Measuring the Search Tactics of Novice and Experienced Searchers, *Journal of the American Society for Information Science*, 44, p.161-174.

Hwang, G.-J. and Kuo, F.-R. (2011) An information-summarising instruction strategy for improving the web-based problem solving abilities of students. *Australasian Journal of Educational Technology*, 27(2), p.290-306.

Iivonen, M. and Sonnenwald, D. H. (1998) From Translation to Navigation of Different Discourses: A Model of Search Term Selection during the Pre-Online Stage of the Search Process. *Journal of the American Society for Information Science*, 49(4), p.312-326.

Ingwersen, P. (1992) Information retrieval interaction. London: Taylor Graham.

Ingwersen, P. (1996) Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation*, 52(1), p.3-50.

Ingwersen, P. (1999) Cognitive Information Retrieval. *Annual Review of Information Science and Technology (ARIST)*, 34, p.3-52.

Ingwersen, P. et al. (2010) Panel: Perspectives on adaptivity in information retrieval interaction (PAIRI). *Proceedings of the 73rd ASIS&T Annual Meeting*, 47, p.1-5.

Ingwersen, P. and Järvelin, K. (2005) Information Retrieval in Context – IRiX. *ACM SIGIR Forum*, 39(2), p.31-39.

Jansen, B.J., et al (1999) The use of relevance feedback on the Web: Implications for Web IR system design. *Proceedings of WebNet World Conference on the WWW and Internet*, p.550-555.

Jansen, B.J., et al (2000), Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing and Management*, 36(2), p. 207-227.

Jansen, B.J. and Pooch, U. (2001) Web user studies: a review and framework for future work. *Journal of the American Society for Information Science and Technology*, 52(3), p. 235-246.

Jansen, B. J. and Spink, A. (2003) An analysis of Web documents retrieved and viewed. *The 4th International Conference on Internet Computing*, p.65-69.

Jenkins, C., Corritore, C.L., & Wiedenbeck, S. (2003) Patterns of information seeking on the Web: a qualitative study of domain expertise and Web expertise. *IT & Society*, 1(3), p.64-89.

Joho, H. (2009) Cognitive Effects in Information Seeking and Retrieval. *ECIR 2009 Workshop on Contextual Information Access, Seeking and Retrieval Evaluation*, Available at: http://www.dcs.gla.ac.uk/~hideo/pub/cirse09/cirse09.pdf [Accessed 20th September 2010]

Karunaratne, S. and Dias W.P.S. (1997). University of Moratuwa: History 1972-1997. University of Moratuwa, Moratuwa.

Kobayashi, M. and Takeda, K. (2000) Information retrieval on the Web. *ACM Computing Surveys*, 32(2), p. 144-173.

Lagoze, C. (1997) From static to dynamic surrogates: resource discovery in the digital age. *D-Lib Magazine* [Online], Available: http://www.dlib.org/dlib/June97/061lagoze.html [Accessed 2nd Aug 2009].

Luhn, H. P. (1957) A statistical approach to mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, 1, p.309-317.

Malik, Amara and Mahmood, Khalid (2009) Web search behaviour of university students: a case study at University of the Punjab, *Webology*, 6(2), p.1-14.

Mansourian, Yazdan (2004) Similarities and differences between web search procedure and searching in the pre-web information retrieval systems. [Online], Available from: http://www.webology.ir/2004/v1n1/a3.html [Accessed 3rd March 2009].

Mizzaro, S. (1997) Relevance: the whole history. *Journal of the American Society for information Science* 48(9), p.810-832.

Navarro-Prieto, R., Scaife, M., and Rogers, Y. (1999) Cognitive strategies in Web searching. Proceeding of the 5th Conference on Human Factors & the Web. [Online],

Available: http://zing.ncsl.nist.gov/hfWeb/proceedings/navarro-prieto/. [Accessed 7th Aug 2009].

Nicholas, D., Huntington, P., Jamali, H. R., Rowlands, I., and Fieldhouse, M., (2009) Student digital information-seeking behaviour in context, *Journal of Documentation*, 65(1), p.106-132.

Patitungkho, K. and Deshpande, N. J. (2005) Information seeking behaviour of faculty members of Rajabhat University in Bangkok. *Webology* [online] 2(4). Available from: http://www.webology.ir/2005/v2n4/a20.html [Accessed 30th Dec 2010].

Pollock, A. and Hockley, A. (1997) What's wrong with Internet searching. *D-Lib Magazine*, Available: http://www.dlib.org/dlib/march97/bt/03pollock.html. [Accessed 7th Aug 2009].

Saracevic, T. (1996) Modeling Interaction in Information Retrieval (IR): A Review and Proposal. *Proceedings of the ASIS Annual Meeting*, 33, p.3-9.

Saracevic, T. (1997) The stratified model of information retrieval interaction: Extension and applications. *Proceedings of the American Society for Information Science*, 34, p.313-327.

Schneiderman, B. et al. (1997) Clarifying search: A user-interface framework for text searches. *D-Lib Magazine*, January 1997. [Online], Available: http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html [Accessed 2nd Aug 2009].

Silverstein, C. *et al.* (1999) Analysis of a very large web search engine query log. *SIGIR Forum*, 33(1), p. 6-12.

Spink, A., Bateman, J. and Jansen, B. J. (1999) Searching the Web: a survey of EXCITE users. *Internet Research*, 9(2), p.117-128.

Spink, A. (1999) Toward a theoretical framework for information retrieval within an information-seeking context. *Proceedings of the 2nd International Conference on Information Seeking in Context*, p.12–15.

Spink, A., Jansen, B. J., and Ozmultu, H. C., (2000) Use of query reformulation and relevance feedback by Excite users. *Internet Research*, 10(4), p.317- 328.

Spink, A. et al. (2001) Searching the web: the public and their queries. *Journal of the Americam Society for Information Science and Technology*, 52(3), p.226-234.

Spink, A et al. (2002) Information-Seeking and Mediated Searching. Part 1. Theorerical Framework and Research Design. *Journal of the American Society for Information Science and Technology*, 53(9), p.695-703.

Sullivan, D. (2009) Search engine sizes. [Online], Available: http://searchenginewatch.com/reports/sizes.html [Accessed 2nd August 2009]

Sutcliffe A.G. and Ennis, M. (1998) Towards a cognitive theory of information retrieval. Interacting with computers, 10(3), p.321-351.

Sutcliffe, A.G., Ennis, M., and Watkinson, S.J. (2000) Empirical studies of end-user information searching. *Journal of the American Society for Information Science*, 51(13), p. 121-123.

UMLS (1994) UMLS knowledge source. 5th Experimental Edition. Bethesda, MD: National Library of Medicine.

University of Moratuwa (2010) University handbook 2010. University of Moratuwa, Moratuwa

Vakkari, P. (1999) Task complexity, information types, search strategies and relevance: Integrating studies on information seeking and retrieval. p.35-54.

Vakkari, P. (2000) Changes in relevance criteria and problem stages in task performance. *Journal of Documentation,* 56(5), p.540-562.

Vakkari, P., Pennanen, M. and Serola, S. (2003) Changes of search terms and tactics while writing a research proposal: a longitudinal case study, *Information Processing and Management*, 39(3), p. 445-463.

Wheatley, H. B. (2002) What is an Index? A Few Notes on Indexes and Indexers (p.28). London: Longmans, Green & Co.

White, Marilyn Domas, and Iivonen, Mirja (2001) Questions as a factor in Web search strategy. *Information Processing and Management*, 37, p.721-740.

White, R.W., Dumais, S.T., and Teevan, J., (2009) Charactering the influence of domain expertise on web search behavior. *Proceeding of the second ACM international conference on Web search and data mining*, p.132-141.

Wilson, T. D. (2000) Human Information behavior. Informing Sciences, 3(2), p.49-55.

Wilson, T. D. (1999) Models in information behavior research. Journal of Documentation, 55(3), 249-270.

Yates, R. B. and Neto, B. R. (1999) Modern information retrieval. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Yin, Robert K. (1994) *Case study research design and methods*. 2nd edition. New Delhi, SAGE publications Ltd.

Zhang, X., Anghelescu, H., and Yuan, X., (2005) Domain knowledge, search behaviour, and search effectiveness of engineering and science students: an exploratory study, *Information Research*, 10(2) p.217.