The Comparative Study of Web Based
Informational Queries Changing:
A Research from User’s Aspect

A study submitted in partial fulfilment
of the requirements for the degree of
Master of Science in Information Management

at

THE UNIVERSITY OF SHEFFIELD

by

Yili Luo

August 2010
Abstract

Previous study indicated that the Web-based informational queries were changing due to the emerging of the “super sites” in different domains on the Web since the new era of Web 2.0 in the English based Web environment. Compared to Google’s position worldwide, Baidu is more popular in Chinese web search engine market with several Chinese domestic “super sites” such as Sina, Renren and Taobao in various domains, which consisted a different picture compared to English based Web environment.

This study aimed to investigate Chinese web users in University of Sheffield in order to confirm the trend that Web-based informational queries are changing not only happened in the English-based Web environment, but also validate in the Chinese-based Web environment.

The design of the study adopted the mixture of both quantitative and qualitative methodologies, over 1000 of questionnaires were distributed through the University’s email system and 124 of Chinese students returned with valid responses with over 30 questions related to the study about their experiences when searching for Chinese information on the Web.

The results demonstrated the Chinese web users’ contemporary experiences when searching information in different domains on the Web. By comparing the data of this study with previous research, the findings shows that web-based informational queries are changing both in the English and Chinese based Web environment and the reasons of this trend were partly revealed as the emerging of the existences of different “super sites”.

It is concluded that the web-based informational queries are changing indicates a major challenge for the Web search engine industry. Further works should be done in terms of addressing how to improve the effectiveness of search engines for a varied set of tasks to better understand users’ intentions behind their changing information needs from a global perspective.
Acknowledgement

I am grateful for my supervisor, Professor Mark Sanderson, for offering all the guidance and helps throughout the research project and devoting his valuable time week by week.

I also want to acknowledge especially the former undergraduates Tann in Department of Information Studies, who conducted a great study to inspire me to explore the fascinated research area.

Thanks to all the participants of this study for their enthusiasms and my classmates who helped me to improve the survey.

In particular, I would like to express my gratitude to my parents for their love and understanding during all phases of the dissertation.
# Table of Contents

Abstract .................................................................................................................. 1
Acknowledgement ................................................................................................... 2
List of Figures .......................................................................................................... 5
List of Tables .......................................................................................................... 6

**Chapter 1 Introduction** .................................................................................... 7
  1.1 Background of the Research ................................................................. 7
  1.2 Research Objective and Significance .............................................. 9
  1.3 Research problems ............................................................................. 10
  1.4 Dissertation Structure ................................................................... 10

**Chapter 2 Literature Review** ...................................................................... 12
  2.1 Information Retrieval on the Web .................................................... 12
    2.1.1 The users of web search engines ............................................. 13
    2.1.2 Users’ information needs ....................................................... 14
    2.1.3 Web Query Analysis ............................................................. 14
  2.2 The taxonomy of web search .............................................................. 16
  2.3 The evolution of Web Search ............................................................... 18
    2.3.1 Web search from Web 0.0 to Web 2.0 .................................. 18
    2.3.2 The relationship between Web 2.0 and Web search ............. 19
  2.4 The web search engine towards differences cultures ....................... 20
    2.4.1 Cultural differences toward web search engines .................. 20
    2.4.2 The development of Chinese web industry at present ........... 21
  2.5 The future trend of web search engine ............................................... 23

**Chapter 3 Methodology** .............................................................................. 25
  3.1 Design of the study ............................................................................. 25
    3.1.1 Quantitative Research Methodology .................................. 26
    3.1.2 Qualitative Research Methodology ..................................... 27
  3.2 Data Collection ..................................................................................... 29
    3.2.1 Pilot Work .......................................................................... 29
    3.2.2 Constructing Questionnaire ................................................. 29
    3.2.3 Administrating Questionnaire .............................................. 30
Chapter 4 Presentation of results ................................................................. 32

4.1 Initial Questions .................................................................................. 32

4.1.1 Questionnaire and Population ....................................................... 32
4.1.2 Demographics of the respondents .................................................... 33
4.1.3 Experience on Internet of the respondents ....................................... 35

4.2 General searching behaviour when searching for Chinese information ........ 37

4.2.1 Preferences of web search engines .................................................... 37
4.2.2 Preferences of information sources ................................................... 39

4.3 Specific site when information searching .............................................. 41

4.3.1 Factual information ........................................................................ 41
4.3.2 Chinese current information ............................................................. 45
4.3.3 Shopping information .................................................................... 48
4.3.4 Multimedia information .................................................................. 52
4.3.5 Social-related information ................................................................. 54

4.4 Further Suggestions by open questions ................................................. 59

Chapter 5 The comparisons of two study ............................................... 61

5.1 Comparative Statistics ....................................................................... 61

5.1.1 Comparisons between General Web searching Behaviours ............... 62
5.1.2 Comparisons between specific types of information ......................... 63

5.2 Implications toward Web Search ......................................................... 64

Chapter 6 Conclusion .............................................................................. 66

6.1 Conclusions ......................................................................................... 66
6.2 Limitations and Recommendations ....................................................... 67

Bibliography ............................................................................................. 69

Appendix .................................................................................................... 73
List of Figures

Figure 1 ................................................................................................................. 33
Figure 2 ................................................................................................................. 34
Figure 3 ................................................................................................................. 34
Figure 4 ................................................................................................................. 35
Figure 5 ................................................................................................................. 36
Figure 6 ................................................................................................................. 38
Figure 7 ................................................................................................................. 38
Figure 8 ................................................................................................................. 39
Figure 9 ................................................................................................................. 40
Figure 10 ............................................................................................................... 40
Figure 11 ............................................................................................................... 42
Figure 12 ............................................................................................................... 43
Figure 13 ............................................................................................................... 44
Figure 14 ............................................................................................................... 44
Figure 15 ............................................................................................................... 45
Figure 16 ............................................................................................................... 46
Figure 17 ............................................................................................................... 47
Figure 18 ............................................................................................................... 48
Figure 19 ............................................................................................................... 49
Figure 20 ............................................................................................................... 50
Figure 21 ............................................................................................................... 51
Figure 22 ............................................................................................................... 51
Figure 23 ............................................................................................................... 52
Figure 24 ............................................................................................................... 53
Figure 25 ............................................................................................................... 54
Figure 26 ............................................................................................................... 54
Figure 27 ............................................................................................................... 56
Figure 28 ............................................................................................................... 56
Figure 29 ............................................................................................................... 57
Figure 30 ............................................................................................................... 58
List of Tables

Table 1 ......................................................................................................................... 61
Table 2 ......................................................................................................................... 62
Table 3 ......................................................................................................................... 63
Chapter 1 Introduction

1.1 Background of the Research

There was a remarkable study conducted by Broder in 2002 found that there were more possibilities of the intentions when web users searching for information, rather than just informational intentions behind the “information need” which has been viewed by classic information retrieval for a long time as the reason, however, the navigational (the known sites which already in users’ mind) and transactional (the sites where users could perform a certain transaction) intentions were also should not be ignored when researchers trying to understand the users of the World Wide Web.

In the year of 2007, a project conducted by an undergraduate student Tann from the department of Information Studies supervised by Professor Mark Sanderson surveyed how students in University of Sheffield as web users were searching for web information at that moment. While the primary question was aim to ask students what their favourite search engine is, and Google was chosen by the majority of students by no surprises since the dominant position where Google had among the English based web search engine market. Later, when this study went further to investigate more carefully what sources the surveyed students have used for searching various kinds of information and a more complex information seeking picture emerged. A number of domain specific information “super sites” now existing around the World Wide Web, such as Wikipedia towards factual information, IMDB towards entertainment information especially information about movies, Flicker towards image information and YouTube toward multimedia information, etc(Tann, 2007). Generally, the students surveyed were found to often by-pass Google, getting their information directly from these “super sites”. After carefully refinements, this student’s work supervised by Professor Mark Sanderson "Are web based informational queries changing" has been submitted to a journal and been published(Sanderson, 2009).

A major concern has been made here is that as the World Wide Web has become a worldwide phenomenon, the necessity of studying this web searching trend across different cultures and regions was not addressed in this study. Additionally, in terms
of web-based queries changing highly relevant with the nature of language, this trend should be further researched with different language, especially Chinese language.

It is true that Chinese language is the second widely used language both in the world and on the Web. As an old and sophisticated language, there are a variety of features of Chinese language are different from English. Apart from the diversity of words or characters, the most related factor with web search engine should be mentioned here is that Chinese language owns “little sign words compared to English” such as “where” or “and” or “but”; besides, language is also based on a national culture with variety and dynamic features (Foo and Li, 2004).

It is well known that the development of Web industry in China has been growing rapidly in recent years. When the Internet has been introduced and being popular in China since the new century, there were massive and urgent needs with the great demands for high-quality Chinese language based web search engine. Not surprisingly, a large numbers of domestic web search engines have emerged in light of this trend along with giant foreign counterparts such as Yahoo and Google.com. According to the previous observations and official statistics, different from Google’s dominant position among English speaking countries, a different picture of Web search industry as well as Web industry are existing in China where the web search engine market is dominated by a local brand web search engine called “Baidu” (Alexa, 2010) and a great number of websites with regional cultural features and language advantages have occupied the major Chinese web market, just like Wiki, Facebook, YouTube, eBay, IMDB and Flicker, these Chinese local brand such as Renren, Tudou, TaoBao and Sina have come to a great success and become the Chinese based “super sites” among various domains in China (Alexa, 2010).

Consequently, it could be viewed as a totally different web environment in China compared to the English based web environment which surveyed in the Tann’s research. In order to confirm the findings made in Tann’s study and better understand the Chinese web users, it is worthy to conduct a survey aim to Chinese web users and made comparisons between the two different web environments.
1.2 Research Objective and Significance

Having decided the research objective is to conduct a investigation among Chinese users’ experiences compared to the Tann’s to work to validate such a shifting trend based on web search whether is also happening when users searching for Chinese information on the web. At the first, all the related themes should be set out here and what this work could contribute to and develop should be suggested in relation to approve this trend. It is necessary to address here that Tann’s study described the results of a questionnaire examining certain aspects of the web-based information seeking practices of University students based on English. These results showed that Web-based queries generated by search engine can be categorised to one of three categories including “informational”, “navigational” and “transactional” queries.

The Tann’s survey(2009) also indicated that “a large group of queries, which in the past, would have been classified as informational have become at least partially navigational, this change has occurred because of the rise of super websites holding particular types of information such as Wikipedia and IMDB”.

Certain aspects of the web-based information seeking practices of University students were examined, which showed a large group of queries have changing from informational to partially navigational due to the rise of large web sites holding particular types of information, such as Wikipedia and IMDB (Tann and Sanderson, 2009). This research is aim to investigate the trend of web based informational queries changing in several specific Chinese language based information sources, and justify with reasons whether this trend in China with different domain search engine and different "super sites" is similar to the adapted research. Eventually, it would be contributed by comparing the trend of the users' search behaviour, particular in the queries change, on the web between the most wide used language English and Chinese to address the issues from users’ perspective to see which change will bring what implications for the development of web search industry.
1.3 Research problems

Mainly, the fundamental research problem in this project has been identified as to confirm the trend that “are informational queries changing” in both English and Chinese language based on the World Wide Web. In another word, when users searching for information on the Web, the using of web search engines to what extent are becoming navigational or transactional tool. Secondly, the main research problem of this study has been divided into four sub-problems, they are illustrated as following:

1. Whether web based informational queries are changing among Chinese users;
2. To what extent of this trend is happening and what types of query have been shifting from informational queries;
3. For what particular reasons existing behind this trend;
4. Of what implications will be made toward web search engines.

1.4 Dissertation Structure

In this dissertation, the organisation is structured into six main chapters; they are expanded step by step as following:

Chapter one has provided an overview of the background and objective of this study along with the problems should be solved in the process of study.

The related studies were discussed in-depth in light of this research in the second chapter.

Chapter three describes the methodologies have been adopted and how the research was designed; additionally, the process of data collection as a crucial part of this study are also included with the details of how the questionnaire was designed and administrated.

Then the results of survey during this research will be presented in details in chapter four with the illustration of various charts.
Chapter five especially compared the Tann’s work with the findings of this study; three tables were outlined to facilitate these comparisons.

Finally, the chapter six is the conclusion of this project to conclude main points have been made and what further works should be done in the near future related to this study.
Chapter 2 Literature Review

The chapter of literature review discussed the previous researches on five domains related to this study. The literature review was started from a general review of Information Retrieval on the Web, and then followed by addressing the significance of the taxonomy of web search. To keep up with the quickly shifting environment of the World Wide Web, the evolution of web search were discussed to introduce several key milestones with the development of web search as well as to investigate factors which have been changing the web search to a great extent at present. Since this research also addressed the significant issues related to cultural impact on the web search, the literatures of web search engine towards different cultures especially Chinese and English were also reviewed in-depth. At the end of this chapter, the recent works have been done to predict the general trends of web search engines in the near future were discussed from the global perspective.

2.1 Information Retrieval on the Web

The classic model for information retrieval start from the users with information need who will structure queries in written language, then queries will be handed to a specific system which will identify the queries from a series of corpus; at this level, by some certain rules, the queries will be matched with selected documents, further refinement then might be needed to generate new queries through query refinement process (van Rijsbergen, 1979). In general, the web search is different from traditional information retrieval due to its “unprecedented in scale” and “lack of coordination” by its nature and dynamic of backgrounds as well as the changing motives of the users on the web (Manning et al., 2008), it also argued by Martzoukou (2005) that users are always should be placed in the centre in this topic. The following sessions in this part will focus on the Information Retrieval based on the World Wide Web from users’ perspective.
2.1.1 The users of web search engines

It is essential to understand the users of web search at the first place. With the rapid development of Web, the web search users tend to become more and more “normal” (Manning et al., 2008) rather than as professional as information scientists. Ten years ago, experts was considered more proficient in using search engine than novice according to a study conducted by Lazonder (2000) and his colleagues, in which defined users who had engaged in “WWW for less than 10 hours” were novice users and users who had “over 50 hours of WWW experiences” were experts users; however, since the online search experiences have been an indispensable part of web users’ daily life now the majority of web users could be viewed as “expert” without becoming an information scientist in the level of addressing the significance of navigating the Web and the ability to identify the pros and cons of the web search engines in order to locate various kinds of information on the World Wide Web (Martzoukou, 2005). From this point of view, Google as a general web search engine is acting as the most successful role in web search with billions of users at present, which has developed two key strategies to deal with this issue: the first one is always focusing on the relevance, the second part is specifically precision and provide a user experience that is extremely responsive (Dudek et al., 2007). Additionally, increasingly researches have been paid attentions on the user-centred approach to evaluate the usability of web search in terms of users’ experiences and feelings. Therefore, it is true that to understand what users are searching for toward a shifting periods perspective has become the heart of successful web search designing for modern Web search engines with the varied nature of World Wide Web(Rose and Levinson, 2004).

Studies of users’ web searching behaviour followed extensively with the popularity and huge success of web search engines, one was considered as an extraordinary effort was to understand users’ web search behaviour and the reasons behind their behaviours has been conducted by Spink (2002) and her colleagues with the ongoing researches to analyse the query logs of the Excite search engine since the year of 1997, 1999 and 2001, and then their researches were followed by Lenvison (2004) and his colleagues and they concluded that the goals underlying the users of web search engines are “many and varied”. On the other hand, apart from the huge
numbers of works has addressed what users are searching for and how the web searching process worked, there was also a few attempts to investigate why users are searching (Baeza-Yates et al., 2006).

### 2.1.2 Users’ information needs

Users’ information needs is still one of the most crucial factors should be considered to impact on web searching processes, which also has been continuing well studied in the user-centred web searching projects recently (Martzoukou, 2005). For instance, O'Day and Jeffries (1993) proposed the information searching process by summarising common “triggers” and “stop conditions” due to people’s changing information needs of their searching behaviours. Additionally, previous studies also have been addressed the issues that allowed people to modify and refine their queries with evolution of their information needs, which argued that the modelling search should be viewed as a serious of information gathering activities rather than a single shoot. However, there were numerous problems with these so called user-centred information searching studies argued by Wang (1999: 67) due to lack of practical techniques to reflect web users’ real information needs. After a few attempts, it was then argued by Martzoukou (2005) that the attentions have been spotted on information need has remained only in the level of theoretical and there was an urgent need for further studies to focus on web users' information need along with the implications behind those interactions. It also indicated Martzoukou (2005) that the answer could be found by redefining the role of the observed users, for instance, with users choosing and describing their web searching tasks by their own rather than the researches imposing users to “search for, locate, identify, select and process information”.

### 2.1.3 Web Query Analysis

As a matter of fact, a web searching process could be simply summarised as the user to submit some keywords, known as web queries, into Question Bar in each session.
A large-scale study conducted by Spink (2002) showed that the majority of web users preferred conducting few query reformulation along with few search terms; additionally, they were difficult to effectively developed relevant key words and applied Boolean queries and researchers found that the users of web search engines mainly generated four types of queries into Question Bar including Boolean queries, single keyword, natural language questions and request. Truly, Rose and Levinson (2004) stated that if we view web search engine as the world, users’ behaviour would be studied under the only focus upon the stream of queries users generated. Web query as the entrance of what users interacting with the web search engines, the changing and shifting pattern of query has been paid more attentions recently by more researchers. Just as it was then emphasised by Rose and Levinson (2004) that the underlying reasons of users’ search behaviours were actually key to satisfying the users’ information need, if the reasons have changed over time, so we should capture the trend and clues from it. This is also the reason why we should learn about the importance of query’s changing in light of the perspective of what they transformed from one to another. Refer to query analysis on the Web search engine was formulated in the nature of language, which factor should not be ignored during such kinds of survey.
2.2 The taxonomy of web search

When the classic information retrieval model defined the reasons for users searching information was driven by “information need” (van Rijsbergen, 1979), which indicated that the nature of most web queries were informational. The statement above has been with no arguments for a long period; however, coming to the age of World Wide Web the situation has been shifted, which statement was argued by Broder several years later that there were three general categories beyond the simply named as information need based on web search, in which queries could be classified as informational, navigational and transactional (Broder, 2002). Since then, the Broder’s taxonomy of web search has provided with different perspectives to review the information retrieval on the Web.

As defined by Broder (2002), users’ search intentions formed as informational queries is to “acquire some information assumed to be presented on one or more web pages”; Navigational queries, on the other hand, is to “reach a particular site”, which implicated that users are intend to find the information has already in users’ mind which existed in specific websites; Transactional queries refer to users who “perform some web-mediated activities”; the typical example is conducting online shopping activities. Broder also determined the balance between the three types of queries, accordingly, 39% - 49% of queries were informational; 30% - 36% were transactional while 20% - 24.5% were navigational queries (Broder, 2002); however, such statistics has been argued by Rose and Levinson that lacked of solid methodology and accuracy due to the two methods Broder has used to classify queries: “a user survey” and “manual classification of log entries”. Two years later, Broder’s categories then have been enhanced and sub-divided by Rose & Levinson (2004), then a more comprehensive and integrated perspective was proposed by their subcategories which divided informational queries into five more areas (directed, undirected, advise, locate, list), while resources queries replaced transactional queries and navigational queries remained in their categories, however, the frequent was less than Broder’s survey. Accordingly, nearly 40% in total were assigned as non-informational queries; on the other hand, rather than learn or absorb actually meanings, the majority of informational queries related to serve the purpose of
locating subjects such as a product or service. As a matter of fact, Levinson(2004) concluded that “just over 35% of all queries appeared to have the kind of general research goal (questions, undirected requests for information, and advice-seeking) for which traditional information retrieval system was designed”. Levinson and his colleagues’ findings clearly stated that the number of navigational queries were over looked existing on the web search and much less prevalent than previous’ study, which was difference from what Broder’s assumption suggested that web search engine should not be encouraged to provide more focus on performance on navigational queries. At the end, Levinson stressed the importance to understand of search goals could provide foundation for future improvement of web search engines.

Then following in 2010, another research investigated a methodology to automatically classify Web queries by topic and user’s intent with more than 20,000 Web query data set (Jansen and Booth, 2010). In this study, they reviewed the previous works with Web query classification and summarised there are five main challenges in this area at present, the first one is the highly dynamic nature of Web content, with nearly everyone with an Internet connection being able to publish on the Web since the era of Web 2.0; the second challenge is the shifting population of user interests which are continuing changing from time to time; the third one is the long tail of user interests and queries, with some highly popular queries and some unusually rare ones; the fourth one is the scale of users, queries and content, “with millions of users, billions of queries per week and even more billions of pages of content”, and this is also will be a growing trend in the near future; the last challenging issue is the nature of Web queries, which is as short as two or three terms and “there are very few inherent attributes” just make the query classification will never be an simple task.
2.3 The evolution of Web Search

Recently, with the transformation of Web, what changes have been made to previous research could be a breakthrough point to study the current Web form web search engines. When first appeared in the early of 1990s, a great number of incremental changes have been consistently made to web search engines; the most successful change brought by Google and its counterparts was the transformation from a classic Information Retrieval based algorithms only supported with informational queries, and then became the second generation web search services supporting navigational queries using advanced web specific information especially link analysis (Broder, 2006). At the same time, web search also has been from an unknown concept became one of the most successful industry in the era of World Wide Web. The initial goals of web search engines hoped to become bigger and faster(Raman, 2009) have been achieved in the second generation web search engines just like Google did. Coming to the era of Web 2.0, how to make a better web search engine has been recently discussed more and more often, and the better means to beyond functionality like Web utilities or Web Resource Centres, but to make better users’ experiences by directly satisfy the users’ intent (Rose and Levinson, 2004).

2.3.1 Web search from Web 0.0 to Web 2.0

Originally, the Web search in the era of Web 0.0 was simply known as an HTML interface in front of a database management system (DBMS), which stressed more on the functions from the perspective of technology rather than how these functions integrated together to work in a better, more powerful way (Hock, 2007). Nowadays, the Web 0.0 has been buried in the tide of history. O'Reilly (2007) has made a good example of how to identify “Web 2.0” from “Web 1.0”, who has tried to extract seven general principles described as “services”, “control over unique”, “trusting users as co-developers”, “harnessing collective intelligence”, “leveraging the long tail through customer self-service” and “software above the level of a single device”; finally, as “lightweight user interfaces”, “development models and business models”.

18
To summarise, the web has been consistently evolving as an incredible online platform allowing users to participate with increasingly activities. Just like what O'Reilly (2007) concluded in the end, “Like many important concepts, Web 2.0 does not have a hard boundary but rather, a gravitational core.” Unlike Netscape or other early attempts that tried to sell information as a product on the web, web search engines such as Google have followed a unique principle that could be seen as not just a collection of software tools, but a specialized database with data as free tools for the general public. Actually, not only Google, but also other successful web search engines, the web search engine in the new era of Web were predicted by O'Reilly (2007) as “in the space between browser and search engine and destination content serve” as well as “an enabler or middleman” between users and the vast World Wide Web.

2.3.2 The relationship between Web 2.0 and Web search

As the rapid growth of Web 2.0 services on the web, it was spotted by Tann and Sanderson (2009) that there has been a noticeable trend referred to a small number of super sites has been accumulated for increasingly large numbers of English information related to some certain domains such as Entertainment information, Factual information and Technology information; Correspondingly, websites such as the Internet Movie Database (IMDB) for entertainment information mainly about movies and television and Wikipedia for general knowledge in an overview pattern as well as other well known sites adopted with Web 2.0 principles and features have been dominant more and more areas on the World Wide Web. Accordingly, this hypothesis was later proved by his survey among university’s students with reliable statistics based on English information; however, such survey intend to reveal the trend on the web has not addressed the fundamental issue was that the World Wide Web has been thrived among other language or culture.
2.4 The web search engine towards differences cultures

Due to linguistic and cultures, it has been generally believed that the World Wide Web also is dynamic with different web structure and layout as well as content. There are also significant differences between English sites and Chinese sites on the Web. Recent studies have been examined that the differences in web characteristics between English sites and Chinese sites and also had certain impact on the users’ experiences of using web search engines(Li et al., 2006).

As the one of the most common language in the world, Chinese has emerged as a key language on the World Wide Web(CNNIC, 2009). Each language owns their unique features such as character and grammar as well as the language based culture, preliminary works have shown that users’ behaviour on the web were influenced by national culture in different level (Liao et al., 2010). It is argued by people that language as a symbol and tool to represent national cultural also has a significant implications on users’ behaviour web. There also has more and more studies focused on understanding users’ behaviours toward different cultures. More specified, Chinese are taught to respect hierarchy and be aware of one’s social status in such a hierarchy, moreover, Chinese culture stresses inter-dependence and group solidarity and harmony, such hierarchical and collectivistic characteristics of Chinese national culture.

Therefore, it was easier to understand the observation that Chinese web has certain characteristic in their structure, layout and content that are much different than English web sites.

2.4.1 Cultural differences toward web search engines

Previous research has suggested that individuals with different national cultural values may tend to respond differently to websites within some specific domains, besides, technology-related beliefs and behaviours also have shown vary with different national cultural values(Li et al., 2009).
Among various studies in the national culture areas, Hofstede’s work has been considered as a classical model, who defined culture as “the collective programming of the mind which distinguishes the members of one group or category of people from another” (Hofstede, 1980). He also proposed five cultural dimensions, include Power Distance, Individualism/Collectivism, Masculinity/Femininity, Uncertainty Avoidance and Time Orientation, these dimensions were developed in general workplace, family and school relationships to reflect a distinct cultural value. Another widely applied model for investigating the intentions to use a technology here should be mentioned here is technology acceptance model (TAM), in which group members’ beliefs and behaviours will be determined by the core values of a culture shaped as the cognitive processes (Li et al., 2009).

Web search engine, as the leading role in the web technology industry, which provides increased support for personalization of contents, language, interaction in the age of Web 2.0, cultural issues has increasingly got attention to a new level. While cultural factors has been widely considered as a main challenges for making progress to design websites that compensate for different culture (Li et al., 2009).

2.4.2 The development of Chinese web industry at present

Unlike the other English-speaking region and most of other area of the world, Google is not the dominant search engine in China. Chinese users prefer the national-brand search engine called Baidu which more powerful in local language and with some popular features favoured among Chinese users; Baidu offers 57 search and community services, such as Baidu Encyclopaedia, an online collaboratively-built encyclopaedia, and a searchable keyword-based discussion forum (Analyst, 2010), they are the key service to attract users in China. Since established in 2000, Baidu now is the No. 1 search engine in China who controlling 63 percent of China's market share (Alexa, 2010). The importance of addressing cultural difference from the users’ experience has been studied from both research and in practice as well. Also according to the latest official report (CNNIC, 2009), Chinese web users are serving by over 300 web search engines, navigating a web space of over 680,000 websites ending with.cn in their URLs. The most popular search engines include
Baidu, Sousou, Zhongsou and Yahoo Yisou. On the other hand, Chinese web characteristic of low updating frequency, high number of internal links in a web cluster etc.

After the huge successes of these user-centred English-based super sites, a similar transcript also has been published in a short period by their Chinese counterparts. Almost in every domain of the online services, a few Chinese-based super sites established and obtained huge market place among the Chinese users on World Wide Web (Li et al., 2006). Just like what Wikipedia, IMDB, eBay and Youtube did, these Chinese websites absorbed their English counterparts advantages and also, as a local brand, following the local culture and adapting to local environment to create massive popular sites welcomed by millions of Chinese users, such as Renren (Former name as Xiaonei, means within the university) attract over 8 million Chinese university students to take part in (Liao et al., 2010); Taobao, which has been the top choice of all online transaction platform among Chinese users; Also happened with Youku and Tudou, which websites are similar with Youtube, and greatly successful among Chinese web users (Li et al., 2009). It was true that English web is much larger and has existed longer than the web in any other language and the world leading web search engines have spend a decade building highly specialised ranking functions for English web pages, as a result, the most effective ranking features such as PageRank, domain taxonomies and click-through information were designed based on English but altered or absent in other languages (Foo and Li, 2004). At the same time, a significant portion of non-English queries have unambiguous translations into English which also function as good English queries.
2.5 The future trend of web search engine

Several years before, in a study which has been seemed as the milestone of the web search domain Brin and Page recognized that some relevant search result are generally more valued by users than others. In their research, they presented a highly scalable, efficient implementation of ranking method which bring the Google great success in the web industry and bring the users with high speed query-based summaries, spam rejection, source clustering, and context-sensitive search (Brin and Page, 1998), which had a significant impact on people’s everyday life.

Since then, researchers have been continuing attempted to build perfect search engine which could fully specify users’ information need by focusing on improving keyword search such as meta-data, natural language and cross-language retrieval (Teevan et al., 2004).

Google is not alone to should the responsibilities to develop the perfect web search engine, although the latest statistical estimated over two-thirds of English based web search engine market occupied by Google with only 20% left to Yahoo, the rest pieces of cake was in a cut-throat competition among numerous small fish; nevertheless, there is still enough sweet for “someone with a smart and fresh approach to come in and siphon off market share from Google” (Maxymuk, 2009:89).

By coming up with a fresh idea with solid implementations, challenges are still could be undertaken, many intelligent approaches such as automatically to learn users’ behaviour over time and to create intelligent tools to crawl hidden information beyond what the current web search engines could be indexed.

However, it is also recognised by researchers that the perfect engine may be not enough since users can not always specify clearly about their information need when the search engine is based on keyword. Especially with the rapid development of Web 2.0, there has been a trend towards information of certain forms, mainly called User-Generated-Content websites, coalescing into a small number of large websites (O’Reilly, 2007). As summarized by Tann, as people’s opinion tend to more reliable than just a paragraph of description, and eye farseeing “super sites” has offered the platform for people to distributing their own opinions, the users of web
search engine may tend to treat the search engine just like a “by-pass tool” to get the “known item” when they searching for specific information when the queries were seemed as informational, which implicates the informational queries may be changed as navigational queries or transactional queries (Tann and Sanderson, 2009). Therefore, for the navigational and transactional queries, the search engines would be less functional and users tend to less care about the features of search engine, just get to the point will finalize the search engine’s work. This trend will absolutely have impact on the current situation of search engine, from both functional perspective and market perspective.
Chapter 3 Methodology

3.1 Design of the study

In this chapter, the methodologies adapted to design this research and how the primary data was collected will be reported. In general, the overall strategy of methodology applied in this study was a mixture of quantitative and qualitative research methods; besides, there were several elementary factors impacted on the process of designing the survey, which will be also illustrated stage by stage.

As the previous work conducted by Tann has already addressed the basic problem "Are web based informational queries changing" (Tann and Sanderson, 2009) with data collected from English-based Web environment; in this research, the objective is mainly about exploring this thoughts either generalisability or representativeness with the comparable data from Chinese users’ web searching experiences; therefore, it was necessary to combine both quantitative and qualitative approaches to investigate the first hind data and then find out what implications behind these phenomena.

Banwell and Coulson(2004) explained the necessity as the researchers “need both quantitative and qualitative approaches to be combined to produce both the holistic view and the robust data needed to triangulate and thereby validate data collected”, which kind of strategy was also has been applied widely recently among web searching studies in order to recognise the importance of fundamental insights from a holistic picture of web search.

Additionally, it is also shown by Martzoukou(2005) that researchers who interested in studying web search behaviours have started to paid more attentions to the “hybrid approaches with an attempt to surpass the rigid qualitative versus quantitative paradigm polarization”.
3.1.1 Quantitative Research Methodology

In the first place, quantitative approach as a logical and less challenging approach to gain useful information from people (Wisker, 2001) was used to carry out the questionnaire to collect the primary data towards the experiences of web searching practises and preferences among Chinese students in the University of Sheffield.

Fundamentally, quantitative research method mainly includes the following terms: Scientific, Objective, Positivist, Reductionist, Contrived, Experimental and Deductive. The quantitative research method could be understood from the descriptions in the book written by Leedy and Dellinger (2005) that “To behold is to look beyond the fact; to observe, to beyond the observation. Look at the world of people, and you will be overwhelmed by what you see”.

Consequently, after the estimations of time, scope and other resources should be needed to complete this study and the expectations based on the objectives of this study, the form of survey was confirmed as using the theory of quantitative research method to design the questionnaire to distribute to Chinese students in the University of Sheffield.

Furthermore, quantitative research methodology is also adapted to response the problems in relation to the “measured variables with the purpose of explaining, predicting and controlling phenomena” (Leedy and Dellinger, 2005), which is also sometimes named the experimental, traditional or positivist approach. In this study, the quantitative research methodology facilitated me to observe Chinese students who using web search engine for seeking various kinds of information with insight in the purpose of validating the trend of web based queries shifting in a different culture and language environment compared to the English based environment. Specifically, the quantitative research method was applied to survey Chinese students with questionnaire to analysis the data and offers the reasons behind the results as well as taking into account the implications of this result.

According to the recent studies, the researches of web information searching behaviours has suffered from in consistencies in terms of method and lacked homogeneity in research foci since the unique area had been drawn massive
attentions by researchers (Martzoukou, 2005). The combinations of qualitative and quantitative methods are expected to generate more comprehensive insights in this area, nevertheless, many attempts has been failed due to various reasons, for instance, many researches were not able to ensure enough samples from web search log to represent general pattern of users’ vary intents (Rose and Levinson, 2004), which could be argued as problematic and a less developed consistent methodological framework of research. Additionally, the real information needs and what elements behind the needs in terms of cultural and social has been an extensively challenging topic to research. When addressed these issues and get to know what priorities should be made in the research, the research methodology could be designed more effectively.

3.1.2 Qualitative Research Methodology

On the other hand, in case the data would not answer the underlying meanings and beliefs of users’ experiences when searching information on the Web, qualitative approach was also then applied. Compared to quantitative research methodology, the qualitative research including the following key terms, they are subjective, Anti positivist, holistic, descriptive, phenomenological and inductive (Hancock, 1998). Qualitative research was used also more broadly in social science studies, for this study, which is paid more special interests to seeking the significance of what people’s experience and their attitudes related to these experiences. This approach also required researcher to look beyond the results but to make more effort to come up with the report containing detailed description, analytical opinion and comparable summarisation.

This combination was proved as an effectively and comprehensively attempt to understand the research problems about web searching activities in more depth due to the nature of qualitative research methodology was described by Strauss and Corbin (1990) as “any kind of research that produce findings not arrived at by means of statistical procedures or other means of quantification”. It is true that by comparing the two data sets in this study and the Tann’s study, the results showed
more information to make sense of the meanings what the users’ valuable experiences when searching different kinds of information on the Web.

Wilson (1981:10) stated user “is not merely driven to seek information for cognitive ends but as living and working in social settings, which create their own motivations to seek information”.

It indicates researchers should be noticed when using qualitative research method to study web users’ information seeking behaviours, various factors should not be neglected in the process of analysing after the data was collected in the user-centred study in terms of information searching behaviours, especially, such research should emphasise the importance of adopting holistic approaches as well as recognise the nature of multidimensional during the information searching process in order to uncover the dynamic reasons behind the surface of these web searching experiences.

In terms of the presentations of results and findings, qualitative research methodology also provided with detailed and combination approach to applying more dynamic language rather than only statistical data to express the dynamic thoughts and vivid pictures related to the special questions and assumptions(Leedy and Dellinger, 2005).
3.2 Data Collection

The university’s email system provided with convenient service for the researchers to distributing the questionnaire. During this study, it was true that with the helpful service for distributing the questionnaire to the target group was an encouraging experience. The Chinese students studying at University of Sheffield in different level has received the questionnaire and many of them from undergraduate students to postgraduate students have paid attentions to finish this questionnaire and which were then proved that their responses were with relatively high validity and reliability when the data were collected and analysed. In the following parts of this chapter, the process from constructing questionnaire to the strategies applied to administrate questionnaire as well as the limitation of this survey will be reported.

3.2.1 Pilot Work

The process of designing questionnaire has experienced from creating to adapting and then developing to maturity after many pilot works. It was worth to mention that although there was limited time to conduct a thorough pilot test beforehand, with the great helps of several classmates from Information School of University of Sheffield, many mistakes and unreasonable settings has been adjusted, which was then proved effectively to address some crucial problems in the original version of the questionnaire, just as Leedy stated that “a brief pilot study is an excellent way to determine the feasibility of your study”(Leedy and Dellinger, 2005), although it took some time initially, it saved the total time ultimately.

3.2.2 Constructing Questionnaire

By the nature of web information search, empirical studies has been seen as a comprehensive way to investigate web search activities recently in order to understand the importance of analysing fundamental differences of users. In light of this exploration of the information searching behaviours on the web, questionnaire has been recognised as an effective approach to collect the primary data from users’
real need without reservation due to the anonymous and random sample to target group. In this research, the target group has been identified as Chinese students who studying in the University of Sheffield.

As Banwell and Coulson (2004) emphasised, through the ‘use of multiple techniques of investigation a range of datasets are being built up’, which can allow ‘analysis to be carried out of any gap which may exist between the expectations of service and the reality of that service as seen through user’s eyes’. In terms of the designing of questionnaire, it will be structured as two parts: the first one are closed ended questions, the second one are the open ended question since most questionnaire are formed by the combination of two types of questions (Moore, 1984).

In this study, in order to reduce the ambiguous answers, the closed ended questions are more widely used since the reason that the easier the questions to answer, the more questionnaires could be finished and collected. However, the data would be less meaningful and restricted by the standard answers if the questionnaire is only presented in a simple and straightforward style. Consequently, the open ended questions was also designed with fewer amounts than the planned expectation in order to gain the opinions from the users’ own words, which was proved to facilitate the research to explain complex reasons behind their standard answers with more valuable information.

3.2.3 Administrating Questionnaire

There was other indispensable criteria has been considered to design a reliable and effective questionnaire. In this study, since the primary goal of this project is to gather as many as users’ experiences to analysis and make comparison with those users who mainly use English web search. As a consequence, relevant high rate feedback and reliable response were seemed as the highest ranked factor to design the questionnaire. To achieve these objectives, several strategies have been adopted in the designing stage of the questionnaire. These strategies will be described stage by stage in the process of designing each part of the questionnaire. For instance, since the survey was mainly about users’ web search experiences in Chinese language, it was also reasonable using Chinese language to design the Questionnaire.
It is assumed that would be better understood by participants with those familiar names of specific websites and better described their intentions and understandings in this pattern. Secondly, proposing open questions has been considered as a low efficient approach in the mail questionnaire distributing process, to stick with the objectives of the study, the amount of open questions has to be cut in a certain pattern. Instead, more scaled questions and click boxes were applied to address some more complex issues. Considering the accuracy to best describe users’ intentions, every word have been refined at a clear and general level after conducting the pilot tests based on several constrictive advices.

Another example could be mentioned here is in the beginning of the questionnaire, several general questions were asked about respondents’ web experiences since previous researches suggested that changes in web searching behaviours occur as a result of increased experience of using the web. this notion led to the study of the behavioural features of users who have different levels of Web expertise. When constructing questionnaire, the questions were addressed about the time spent on the web and frequency of using web search engines by students as required answers to identify whether they could be seen as beyond novice of Web users. The assumptions have made by previous studies suggested that more experience users have acquired the more efficient, sophistic and systematic their information seeking strategies and tactics.

As explained by Holscher and Strube(2000) that ‘the knowledge and skills necessary to utilize the WWW and other Internet resources successfully to solve information problems’, in this case, those who could be categorised into experienced users could contribute the data and experience more reliable and useful by possess the ability to recognise the functionality and the limitations of the Web as tool for locating information. There were also other considerations behind each question and they will be further introduced in the next chapter.

After data collection, it was also worthy mention that another underlying reason for the combination of both quantitative and qualitative approaches was due to the research was confronted with a series limitations which has been expected in the planning phase of the research, such as the response rate was relatively low since several reasons. These limitations will be further described in the next chapter.
Chapter 4 Presentation of results

In this Chapter, the results of the survey throughout the whole study of the
documents experiences of web searching practises and preferences among Chinese students in
the University of Sheffield will be represented.

Generally, the survey in the form of questionnaire distributed by the University’s
electronic mailed system was divided into four main parts. The results will be
presented following the original order of the questionnaire.

4.1 Initial Questions

This questionnaire was designed by Google document system and then distributed in
the middle of July during the summer holiday via University’s email system.
According to the latest report by the University, there are 1,058 Chinese students
studying in the University of Sheffield for different course. Therefore, a total
audience of approximate 1058 Chinese students who are currently studying in the
University of Sheffield were invited to fill the online questionnaire.

4.1.1 Questionnaire and Population

According to the final results of the closure date one month after the distributing date,
out of all 1058 Chinese students there were 124 (11.7%) students filled in valid
responses with this questionnaire in all 26 essential questions, which is a highly
encouraged response rates since the time period was during the summer holiday.
Additionally, many students contacted me after filling the questionnaire for further
suggestions about this topic, which brought many unexpected suggestions and which
will be discussed in the final session of this Chapter.
4.1.2 Demographics of the respondents

The first part of the questionnaire was included with five initial questions in order to gather the basic information of the participants. Except the intentions behind the questions about their web experiences which described in the Methodology Chapter, the other questions will be explained one by one in this part.

Firstly, it was interesting to see that there were both 62 females and males out of the 124 responses filled in the valid responses. In this case, the equal percentage (50%) of male and female students could be viewed as no distinguishing in the aspect of gender, which would eliminate the bias form the study to a certain degree (Question 1).

Out of the 124 responses, the majority of the students 69 (56%) filled with valid questionnaire are aged among 23-25 group, followed by 40 (32%) of respondents are within the 18-22 years old bracket, while there were only 10 (8%) and 5 (4%) are aged between 26-30 and over 30 respectively. The reason for setting these age groups is because of the target group is all students who are currently studying in the University, it is also reasonable to believe young people are generally more active and enthusiasm toward the Internet (Question 2).
Figure 2.

Referred to the education level, 21 (17%) are currently taking the Pre-Undergraduate Course; similarly, 15 (12%) are currently taking Undergraduate Course; 22 (18%) of respondents answered that they are currently taking Pre-Postgraduate Course.

The number of Chinese students are currently taking Postgraduate Course is the most among all groups, there are 56 (45%) choose this option, which results may explained the majority of students 69(56%) are aged among the group of 23-25, and this group is considered that they provided the most valuable information about their web searching experiences due to they are relevantly in a higher education level than undergraduate students and less stressful than the students who are taking the PhD’s course. Finally, there were only 4 (3%) are currently take PhD Course (Question 3).

Figure 3.
4.1.3 Experience on Internet of the respondents

As mentioned in the previous chapter, to ask the respondents about their web experiences since previous researches suggested that changes in web searching behaviours occur as a result of increased experience of using the web. Consequently, when asking about how much time has spent on the World Wide Web per week, as much as 62 (50%) of people choose the spent more than 30 hours on the Internet; on the contrary, only 3 (2%) responded that they only spent less than 10 hours on the Internet per week; 22 (18%) and 37 (30%) of the participants specified they spend 10-20 hours and 20-30 hours on the Internet, which results indicated that most of students could be viewed as experienced web users and they are familiar with the World Wide Web since they are using their daily life (Question 4).

![Pie chart showing web usage frequency](image)

Figure 4.

In order to introduce the next part of the whole questionnaire, the questions “how often do you use web search engine searching Chinese language information” were asked in the last questions in this part. Over half of the responses 73 (59%) clarified they do it frequently everyday; only 8 (6%) of the responses answered their search experiences about Chinese information are less than 5 times per week; the others, 33(27%) and 10 (8%) respectively also confirmed that they also used the search engine access Chinese web information everyday from 1-3 and 4-6 times respectively. The results validated the assumption made in the previous questions that the majority of the responses could be viewed as experience web users and at
least non-novice of web search engine users, which would be benefit for the further parts of this survey to a great extent (Question 5).

Figure 5.
4.2 General searching behaviour when searching for Chinese information

In the second part of the whole questionnaire, the general searching behaviour of students in relation to their preferences of using web search engine for searching Chinese information on the World Wide Web were surveyed. In order to gain more information about their general web experiences, the preferences of which kind of information sources they will choose were divided into two categories as local and foreign brands as well as their reasons of choice were also investigated. Generally, there were five questions in this part associated with general searching experiences for target users. The results were shown in the two parts as below.

4.2.1 Preferences of web search engines

To begin with, followed by the last question in the previous part of the questionnaire, the students were asked in the first question to tick their preferences of Chinese information on the web among various areas including “factual information”, “current information”, “academic information” and “multimedia information”, “shopping information” and finally “social-related information”. It was shown that current information draws the most attentions, out of 124 respondents there are 98 (79%) ticked this option; the current information was followed by academic information with 85 (69%) of students choose it as their interesting area, which was a reasonable choice since all of the participants are college students. It was similarly that the factual information as popular as multimedia information by the choice of 76 (61%) and 68 (55%) of students; however, shopping information and social-related information got less attentions by responds with only 38 (31%) and 35 (28%). In fact, this question set here was intention to investigate to what extent the students are using the various kinds of information on the web, and the results showed that all the listed information are essential parts in their web experiences (Question 6).
In regards of which web search engine in general is the first choice when students are searching for Chinese information on the web listed in the answers, Baidu hold a safe lead among respondents by 20% with up to 73 (59%), Google.com was chosen by 48 (39%) respondents as their preferred web search engine. On the other hand, the other web search engines listed here hardly got attentions by respondents: Sousou as a recently lunched local brand web search engine only got 1 (less than 1%) students’ choice; the Chinese version of Yahoo web search engine Yahoo Yisou with 2 (less than 2%); surprisingly, Bing as the integrated web search engine with Windows Operation System, no respondent preferred to using Bing to search Chinese information (Question 7).

The followed question asking students to tick the reasons why choose the above web search engine as their first choice in general to searching Chinese information on the Internet. Based on the standards of evaluating web search engine from users’
perspective, there were six reasons in total listed for respondents to choose (Li et al., 2006). The majority of people 90 (73%) ticked “easy to use” as the main reason for their choice of web search engine, the following reasons were chosen by almost as many respondents as each other range from 26% to 41%. The results were shown as “More updated information” with 32 (26%), Fast speed of the searching process” with 44 (35%), “More reliable” with 39 (31%), “High-quality search results in the ranking system” with 49 (40%), “No language barrier” with 51 (41%). An “other” box also listed to let users to specify their unique reasons showed that 28 (23%) believed there are more underlying reasons for their choices.

![Figure 8](image_url)

**Figure 8.**

**4.2.2 Preferences of information sources**

The question “for Chinese language information websites, which categories of source would you prefer” was asked as an essential preference choice. “Domestic Brand” and Foreign Brand” were the only two categories in this question; the objective for these questions is to know if the respondents hold the ability to tell which kind of information sources they are using. According to the results, it is interesting to see that the domestic brand was lead by nearly 50% from foreign brand with 92 (74%) students’ favour. The example listed here was the most popular sites based on the official statistics (CNNIC, 2009) to illustrate the Chinese websites and the Chinese version of their English counterparts. The comparison result clearly showed that it is true that Chinese domestic brand websites are more popular among Chinese users even the respondents are studying in an English-based environment, which is in turn proved the pre-conditions proposed for this study (Question 9).
The last question was to ask respondents to specify their choice by the listed reasons. There were slight changes compared to the questions to ask respondents to specify their web search engine choice due to the reason of there was a more general aspect should be considered such as cultural factors and reading habits. According to the results, “Easy to use” with 69(56%), which was not the mainly reason to explain their choice; “Fit for my reading habits” accounted for 85(69%) suggested people view this factor in a high level during their information seeking process and implied their information seeking behaviours more involved with browsing activities; “More reliable” with 39(31%), “Faster, just save the time” 28(23%), “No language barrier” and “Various interesting Application integrated in the websites” draws the same selection number with 25(20%), but only 21(17%) of students choose “Great interface” as the reasons for their preference (Question 10).
4.3 Specific site when information searching

The following five different sessions was designed as a similar style to keep a consistence between various kinds of information with four questions in each session, which has been considered to be better understood the key problems related to this study and avoid falling into irrelevant discussions such as the nature of information. Besides, the overall trend will be analysed more convenient in this style. Additionally, every task has been given the appropriate example to make the participants clearly identify the area and help to recall their searching experiences.

The findings will be listed separately in this part by illustrate the certain consistence and inconsistence among all surveyed areas of information. Moreover, some problems such as limitations and the issues will be discussed in the final session of this part.

4.3.1 Factual information

The first question ask students to answer “which is their first choice of information source when they searching for Chinese factual information on the World Wide Web”. The answer has been listed as Baidu (Baike and Zhidao), Goole.com, Wikipedia Chinese Version, Chinese Portal Website (such as Sina, Sohu or 163) and Personal Blog.

The reason to put the “Baidu”, “Baidu Baike” and “Baidu Zhidao” together is because that “Baidu Baike” and “Baidu Zhidao” are actually integrated with the “Baidu” web search engine, when the queries contain factual information such as people and date or related natural language information, the web search services would be automatically navigated to “Baike” or “Zhidao”; additionally, these services could be clicked on the top area of the homepage of Baidu.

According to the results, the majority of respondents 83 (67%) chose a series ranges of services offered by Baidu to conducting their searches for Chinese factual information; Google.com was away behind the top, only 21(17%) responded that they used Google searching for Chinese factual information. In the third place, Wikipedia Chinese Version was as expected took a small pieces in this domain and
was chosen by 16 (13%) respondents, the main reasons could be explained by they were just starting to expand their Chinese language items recently with limited information available; however, Wikipedia Chinese Version was a newly emerging forces and could not be neglected considering their young age in Chinese language. On the other hand, Chinese Portal websites and Personal Blog were not chosen by any students, which showed that they were no longer good factual information sources even they were the main forces in this areas.

![Factual Information](image)

In a follow up question, students were asked to specify by what mainly approach they are taking to access the above source they chose. This is the question used to distinguish between navigational and non-navigational queries when the respondent searching for Chinese factual information. There were five main approaches listed here to let student to recall their most applied experiences. To specified, except “Directly type URL into Address Bar” could be classified as non-navigational approach, all of the other approaches could be seen as navigational approach such as “Type query into search engine to get there”, “From my Bookmark”. The results showed that out of 124 respondents only 20(16%) of students chose directly “type URL into Address Bar”; not surprisingly, apart from approximately half of students used web search itself to seeking factual information without “known sites” in their mind, the others account for 33% of respondents choose the navigational approach to get their first choice of factual information that “type query into search engine to get
there” with 20(16%), “From my Bookmark” with 12(10%), “From navigational websites (examples were given as hao123 or Very CD)” with 4(3%), The last option was “Personal Homepage such as my Baidu or iGoogle”, which also should be categorised into navigational approach since the contents on the page were published by web search engine itself with 7(6%).

![Approach to access](image)

**Figure 12.**

Since there was a dominant number to choose web search services by Baidu and Google.com from last question, it was evident that people tend to start their searching process from web search engine. However, there were several overlap areas due to the blur boundaries between the web search itself and their integrated services, for instance, there were factual service called “Baidu Baike” and natural language service named “Baidu Zhidao” integrated with Baidu web search. *It should be noticed that there was a decrease from 67% to 49% as shown by the results.*

In terms of the reasons why they choose the above source as their first choice, “Accurate and in-depth enough” 56(45%), “More convenient to locate information” 62(50%), “I need high-quality information” 36(29%), “Fit for my reading habits” 52(42%), “Like to participate in editing contents” 11(9%), “More updated information” 28(23%), “Just save time” 54(44%).
In the last question in this session, students were asked their level of experienced of searching for factual information on the web by scale from 1-5, which presented that from unfamiliar to familiar. Accordingly, no one think they find it was very difficult to searching factual information, in the contrary, most respondents believed that they could easily 53 (43%) or very easily 56 (45%) find target factual information by the choosing information sources. This result showed that the majority of students were familiar with the process of finding Chinese factual information with the existing web services.
4.3.2 Chinese current information

As same as the previous session, students were asked four questions about their searching experiences when they looking for Chinese current information, the examples given here shown as the latest hot topic like World Cup news or the World Expo 2010 holding in Shanghai in order to recall their fresh memory in a real set.

The first question is asking students to choose their first choice of information sources regards to Chinese current information. The Chinese Portal News sections were the most popular places to go when the students want to check for Chinese current information chosen by over half of respondents 63 (51%). The others shows a relatively low popularities, Baidu News with 14(11%), Google News Chinese version with 15(12), Social Websites news features (e.g. QQ or Renren) with 20(16%) and RSS News with only 4(3%).

![Current Information Chart](image)

Figure 15.

The reasons were asked in the followed question to investigate why they choose the above source as their first choices from users’ perspective. “Fit for my reading habits” 71(57%) was the most chosen option, since the Chinese portal websites were chosen by most respondents, this reason indicated when web users searching for current news information, they prefer the style to browse the news line by line according to the features of portal websites, besides, these news section such as Sina news offered
“Special Report” section which was popular with web users when they seeking information related with a hot topic indicated by web users in the “other” boxes. From a range of other given reasons, the results showed that “Always find necessary information” was chosen by 51(41%), “Good layout and content” was chosen by 49 (40%), “Save time” with 45(36%), “More reliable” with 27(22%), “I can easily make comment” with 14(11%), “In-depth enough” was chosen by 25(20%). These results also indicated that when web users seeking for current news information they tend to be a passive receiver rather than to directly search for news information due to there are many good existing source they already used to reading current information on them.

![Figure 16](image)

The students were then asked to identify which approach they using to find their first choice of information sources. The results as showed in the following:

“Type query into search engine to get there” with 19(15%)

“From my Bookmark” with 29(23%)

“From navigational websites (examples were given as hao123 or Very CD)” with 12(10%)

“Personal Homepage such as my Baidu or iGoogle” with 13(10%)

“Directly type URL into Address Bar” with 28(23%)

The “other” option was chosen by 23 (19%).
The results of “from my bookmark” and “Directly type URL into Address Bar” were lead by other options indicated that respondents are more familiar with their favourite sources and tend to not using web search engine to get there. It is also worthy mentioned that the increase compared to the “Chinese factual information” in the option of “From navigational websites” and “Personal Homepage” also illustrate the fact that people are clearer about their choice and then by click through the target information source rather than type query into web search engine to get information towards this type.

In the last question in this session, students were asked their level of experienced of searching for Chinese current information on the web by scale from 1-5, which presented that from unfamiliar to familiar. Similarly, only one think they find it was very difficult to searching Chinese current information on the World Wide Web; on the other hand, the other respondents thought that they could easily 46 (37%) or very easily 53 (43%) find target factual information by the choosing information sources, it also should be noticed that there were 12(17%) of respondents choose they cannot be sure their level of familiar with the process. In general, this result showed that the majority of students were happy with the process of finding Chinese current information with the existing web services.
4.3.3 Shopping information

With the prevalence of e-commerce among Chinese young people especially among University’s students, the online shopping information has been continuing attracted increasingly attentions and there would be no accuracy to reflect the student’s real web searching behaviours if exclude the information in this area. The third session in this whole part asked the students about their web searching experiences when seeking for shopping information on the World Wide Web.

The examples also have been given in the beginning of the session introduction; include the most frequently buying or selling commodities such as books, digital products, virtually products and clothes, etc.

The first question was asking students to choose from five options to specify “which is the first choice of information sources related to shopping information”. Five options in the answer area are involved with the most popular online approaches to access shopping information in every domain such as B2C and C2C, also web search engines. Surprisingly, out of 124 responses there are 88(71%) of respondents stated that Taobao as their first choice of shopping information sources, which was an online customer to customer platform offering billions of products by millions of sellers across all Chinese language covered areas. Once the dominant online B2C websites by the official report due to the excellent book selling performance (CNNIC, 2009), DangDang now with only 5(4%) of students considering their information, it
is more like a Chinese version Amazon selling various kinds of products with great discounts recently, however, it was now hanging back and falling behind Taobao, the latecomers greatly surpass the old-timers. With the mass-based users group, BBS still is a very popular place and convenient way to conduct transactions between buyers and sellers, however, in terms of the usability of searching shopping information, which was not a great approach as shown in the results that only 1(%) of students using BBS as their shopping information sources.

Web search engines were also not considered by the majority of students as their first choice that Baidu with only 5(4%) and Google with slightly higher by 11 (9%) of students’ preference.

![Shopping Information](image)

**Figure 19.**

Then the respondents were asked the reasons why they chose the above websites as their primary information source. There are six reasons listed here considering the related features of shopping information from users’ perspective. The result was shown below:

“Convenient to purchase” is the most chosen reason with 64(52%);
Followed by “Comparable price information”, which was chosen by 60(48%);
“Just save time” with 46(37%);
“More reliable, I like to get more information on people’s views” with 45(36%);
“Fit for my shopping habits” with 32(26%).

“Great shopping experiences” with 31(25%);

Out of the 64 (over half of respondents) who chose the option of “convenient to purchase” showed that when Chinese students searching for shopping information they tend to consider the buying the products rather than just to look at the detail or features of this product; additionally, “comparable price information” and “more reliable, like to get more information on people’s views” also prove this information seeking process with the high possibilities to conduct transaction activities rather than just information need about the shopping information.

Then students were asked about their approaches to get the target shopping information sources, the results showed as following:

“Type query into search engine to get there” with 21(17%)
“From my Bookmark” with 32(26%)
“From navigational websites (examples were given as hao123 or Very CD)” with 12(10%)
“Personal Homepage such as my Baidu or iGoogle” with 6(5%)
“Directly type URL into Address Bar” with 33(27%)

The “other” option was chosen by 20 (16%).
In light of the trend showed in the above question, “from my bookmark” and “directly type URL into address Bar” as expected even higher than the percentage showed in current information (23%), which indicated that when respondents searching for information related to shopping they know even better their target information source instead of bypassing web search engine.

In the end of this session, students were asked their level of experienced of searching for Chinese shopping information on the web by scale from 1-5, which presented that from unfamiliar to familiar. Similarly, only 5(4%) and 7(6%) of respondents respectively stated that they find it was very difficult or difficult to searching Chinese shopping information on the World Wide Web. The other respondents thought that they could easily 32 (26%) or very easily 52 (42%) find target shopping information by the choosing information sources, it also should be noticed that there were 27(22%) of respondents choose they cannot be sure their level of familiar with the process.
4.3.4 Multimedia information

In this session students were asked the questions related to multimedia information such as audio or visual information. There were also four questions in this session would be involved to investigate the searching behaviour of respondents when they searching for multimedia information on the World Wide Web.

In the first place, students were asked to choose which source they would prefer when searching for multimedia information. Among all the available web services related to multimedia information in Chinese language, five options were listed to let respondents to choose plus the other option. Accordingly, Out of 124 of the all respondents, 48(39%) chose multimedia sharing websites in particular with Tudou and Youku, which are in a similar style of Youtube in Chinese language; then this result was followed by social sharing websites with 26(21%) of respondents chose this source, it should be considered that when looking the services in these two different kinds of information source, I found that the majority of audio or visual information were embedded with the links from the multimedia sharing websites into the social sharing websites, which is to say, over half of the respondents was obtained their multimedia information from multimedia sharing websites such as Tudou and Youku. On the other hand, the Chinese Portal websites and other sources shared the 40% left spaces with Baidu and Google.com. It was shown that Portal websites with 6(5%), Baidu with 15(12%), Google.com with 8(6%), and 21(17%) chose other sources.

![Multimedia Information](image_url)

**Figure 23.**
As usual, the following question was asked to let respondents to specify their reasons to choose the above sources as their first choice to searching for multimedia information. “The enough collection of multimedia information” was shown as the one of the most vital reasons to impact on respondents’ choice and popular with 81(65%) of students. “Fast enough” and “Update quickly” were with almost equally weighted among the reasons with 57(46%) and 61(49%) respectively. Similarly, out of 124 respondents 46(37%) and 45(46%) respectively of students chose “Easy to locate target information” and “Good quality” as their reasons to choose the source for searching multimedia information.

![Figure 24.](image)

In the third question in this session, students were also asked about the approaches they got to the target shopping information:

“Type query into search engine to get there” with 24(19%)

“From my Bookmark” with 35(28%)

“From navigational websites (examples were given as hao123 or Very CD)” with 7(6%)

“Personal Homepage such as my Baidu or iGoogle” with 5(4%)

“Directly type URL into Address Bar” with 31(25%)

The “other” option was chosen by 22 (18%).
Finally, students were asked their level of experienced searching for multimedia information on the web by scale from 1-5, which presented that from unfamiliar to familiar. Surprisingly, no respondents stated that they find it was very difficult or difficult to searching multimedia information on the World Wide Web. The results showed that out of 124 respondents, 38(31%) and over half with 68(55%) of students believed that they could easily or very easily find multimedia information from the choosing information sources, it also should be noticed that there were 18(15%) of respondents choose they cannot be sure their level of familiar with the process.

4.3.5 Social-related information

The Social-related information has been played an essential part in people’s online experiences, especially among the students in university. This is the session designed
to investigate students’ online searching behaviours towards Social-related information, which could provide a truly requisite piece for this whole study.

Four questions were asked in this session, started with the questions “Which is the first choice of information source when searching for social-related information”. It should be noticed that there was a fierce competition from various competitors based on Chinese online social-related market, to specify, Renren (the former name as Xiaonei, means students within the University) was extremely popular among Chinese students especially among students studying in the University to share their social life online with classmates or friends. The Q-Zone was also an online community integrated with the friend list from QQ, the most successful instant message tool based on Chinese language with more than 200 million users. Douban was developed from a book review website and now became a huge online community with large numbers of sub-groups divided by various kinds of interests, which also called for the group activities in real life based on the large cities in China.

It was also should mention that the latest online social tool named micro blog, among all the sites, Sina micro blog has been the most successful Chinese language based brands, just like the position of twitter in the English based market. According to the results, Renren took a huge leap among all other counterparts with 76(61%), the main reason here should be noticed was that Renren was mainly target at the students, and this survey were also conducted with students. With the large numbers of IM users, the Q-Zone was followed Renren as the second most popular sources when people searching for social-related information. The others were still took a small market place in this area including web search engines like Baidu and Google.com with no more than 6% of respondents chose these sources as their preferred social-related information sources.
The following question was asked to let respondents to specify their reasons to choose the above sources as their first choice to searching for social-related information. “Easy to connect with friend” was shown as necessary for shaping respondents’ choice and popular with 82(66%) of students. “Real and reliable” and “Integrated instant message function” were with almost equally weighted among the reasons with 46(37%) and 45(36%) respectively. Interestingly, “I like their interesting applications” was chosen by 51(41%) of respondents as their reasons to choose the sources. Additionally, out of 124 respondents 38(31%) and 22(18%) respectively of students chose “Fast loaded and update quickly” and “Great interface” as their reasons to choose the source for seeking social-related information.
In terms of the approaches to get their first choice of information related to social-related information, the result was shown as below:

“Type query into search engine to get there” with 17(14%);

“From my Bookmark” with 44(35%);

“From navigational websites” with 9(7%);

“Personal Homepage such as my Baidu or iGoogle” with 9(7%);

“Directly type URL into Address Bar” with 33(27%);

The “other” option was chosen by 12(10%).

![Approach to access](image-url)

**Figure 29.**

Similarly, students were asked their level of experienced of searching for social-related on the web by scale from 1-5, which presented that from unfamiliar to familiar. Similarly, only 2(2%) and 3(2%) of respondents respectively stated that they find it was very difficult or difficult to searching social-related information on the World Wide Web; on the other hand, the other respondents thought that they could easily 39 (31%) find target social-related information by the choosing information sources, and more than half of students with 52 (42%) stated it was very easily to accomplish the searching process; it also worth mention that there were 17(14%) of respondents choose they cannot be sure their level of familiar with the
process. To conclude, this result showed that the majority of students were happy with the process of finding social-related information with the existing web services.

**Figure 30.**
4.4 Further Suggestions by open questions

There are only three open questions integrated in this questionnaire; additionally, all the three open questions were designed not to be essential to answer as well as placed between each part of the questionnaire. The main reasons to do so are because open questions are less interested by respondents in the form of electronic mailed survey which may cause low response rate and placing too many open questions were argued not to be an efficient approach conduct such survey towards web search behaviours since people will not clearly specific about their preferences and underlying reasons in their own words. Consequently, these three open questions were potentially with great value in order to gain respondents’ additional opinions and suggestions.

Specifically, the first open question is located in the end of the second part of this questionnaire. Students were invited to write down their attitudes or suggestions in light of the questions they were asked about their general web searching behaviours. Fortunately, many respondents offered their further suggestions such as there is a trend that many desktop social applications dominant their web searching habit and they do not know exactly what web search engine they are using for searching information; in terms of searching for Chinese information, many of respondents wrote that they were not happy with the accuracy of the searching results among all the web search engine and thought this situation were not changed for 2 or more years.

The second open questions is targeted to obtain further suggestions toward specific information in five domains listed when they are searching them on the web; moreover, they also were invited free to comment on this questionnaire in order to address the missing point should be covered or questions need to improved in the survey. According to the results, students wrote such as some of them use the strategy of typing searching terms plus the target known websites short name as the queries into web search engine and then go straight in the first results in the purpose of convenient; additionally, several respondents suggests they were less in the process of searching shopping information on the Web and this questionnaire should be asked in the first place whether they like to offer their answer in the domain like shopping information.
The last open question was invited students to offer their general suggestions after finishing this questionnaire. However, compared to the previous open questions, this question was obviously less answered and listed some irrelevant suggestions about the objective of this survey, such as there were some region and copyrights restricted when they were searching the multimedia information from Chinese-based websites since they were in the outside of China Mainland.

Last but not the least, it was worthy mention that this questionnaire has been received the most responses in the first day in the 3rd July 2010 which distributed through University’s email system, then it was sharply declined in the number of respondents until the closure date in the 29th July 2010, which indicated that this online survey was only actively and effectively for a short period of time and the time bar came quickly, it could be further improved this situation by repost this questionnaire after a given period of time, for example, after fifteen days.
Chapter 5 The comparisons of two study

In this chapter, the web searching behaviours of English-based survey conducted by Tann in the previous study will be compared with Chinese-based survey in this study. Firstly, two data sets collected from both study will be analysed to find the similarities and differences; then the reasons of their similarities and differences will be discussed along with comparisons in each part of the two surveys; furthermore, what implications brought by the results of comparisons toward web search will be suggested in the end of this chapter.

5.1 Comparative Statistics

In the first table the basic information of two studies will be listed in the following parts:

a. Populations  
b. Response Rate  
c. Questions Numbers  
d. Types of Information involved  
e. Questionnaire Language

Table 1 listed the basic information from the previous English-based study conducted by Tann and the corresponding information in this study.

<table>
<thead>
<tr>
<th>Categories</th>
<th>English-based survey</th>
<th>Chinese-based survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Populations</td>
<td>4,482</td>
<td>1,058</td>
</tr>
<tr>
<td>Valid Responses</td>
<td>220 (5.0%)</td>
<td>124 (11.7%)</td>
</tr>
<tr>
<td>Questions Numbers</td>
<td>45</td>
<td>33</td>
</tr>
<tr>
<td>Types of Information</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Questionnaire Language</td>
<td>English</td>
<td>Chinese</td>
</tr>
</tbody>
</table>

Table 1: English-based survey and Chinese-based survey data sets
There are three points should be explained here according the listed information in the table 1. Firstly, all the respondents in the two surveys are the students who studying in the University of Sheffield during the survey time. The Tann’s survey assumed all the respondents used English as their main language used on the Web and using English questionnaire to conduct the survey; on the other hand, in this study was more focused on investigating students who searching for Chinese information specifically on the Web and using Chinese language questionnaire conduct the survey. Secondly, the reasons for the more numbers of questions involved in Tann’s study is because there were more types of information involved; however, some types of the information could be argued that should be integrated, such as visual information and audio-visual information could be integrated as multimedia information, which approach was applied in this study.

5.1.1 Comparisons between General Web searching Behaviours

There are mainly 10 items could be compared between these two studies in terms of the general Web searching behaviours. The data analysis in this study has the following parts:

<table>
<thead>
<tr>
<th>Categories</th>
<th>English-based survey</th>
<th>Chinese-based survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics of respondents</td>
<td>126 female / 94 male</td>
<td>62 female / 62 male</td>
</tr>
<tr>
<td>Main age group</td>
<td>18-21 (69%)</td>
<td>23-25 (56%)</td>
</tr>
<tr>
<td>Main taking degree</td>
<td>Undergraduate</td>
<td>Postgraduate</td>
</tr>
<tr>
<td>Search engine experience</td>
<td>Everyday (69%)</td>
<td>Everyday (94%)</td>
</tr>
<tr>
<td>Main used search engine</td>
<td>Google (96%)</td>
<td>Baidu (59%)</td>
</tr>
<tr>
<td>Main reasons for above choice</td>
<td>Faster (56%)</td>
<td>Easy to use (73%) and</td>
</tr>
<tr>
<td></td>
<td>And Convenient (56%)</td>
<td>No language barrier (41%)</td>
</tr>
<tr>
<td>Considered variables</td>
<td>Google and Wikipedia</td>
<td>Local websites and foreign websites</td>
</tr>
</tbody>
</table>

Table 2: Comparative statistics of general web searching behaviours for English-based and Chinese-based data sets

According to the table 2, the first point should be noticed is that the age groups between two surveys were slightly different. The English-based survey received the
most responses from age between 18-21 and students who are taking undergraduate course; on the other hand, the Chinese-based survey received the most responses from age between 23-25 and students who are taking postgraduate course. This results suggested that the Chinese-based survey may received more experienced web users’ responses, which then proved by the following items listed that in the Chinese-based survey 94% of respondents answered they are using web search engine everyday compared to the 69% in the English-based survey; additionally, compared to undergraduate students, the postgraduate students might provide more accurate information due to they are more reliable to hold a fixed pattern toward web searching behaviours.

Secondly, the preference of web search engine has proved the previous assumption that when using web search engine searching for Chinese information, Chinese students prefer to using Baidu (59%) as their first choice of web search engine, which is different from the English-based survey showed that the dominant position of Google with 96%. Then, the differences of main reasons for choosing web search engine also suggested that when searching for information in different language, users may take different views, such as English-based survey showed that “faster” is the main reasons for choosing Google as their web search engine; however, it was only selected by 35% of respondents in Chinese-based survey as their reasons for choosing web search engine.

5.1.2 Comparisons between specific types of information

As the previous study has involved seven types of information and has slightly differences in terms of information types, the comparisons will be focused on the similar information types, the results shown as below:

<table>
<thead>
<tr>
<th>Categories</th>
<th>English-based survey</th>
<th>Chinese-based survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factual information</td>
<td>First choice source</td>
<td>Google (62%)</td>
</tr>
<tr>
<td></td>
<td>Main Reasons</td>
<td>Necessary enough (49%)</td>
</tr>
<tr>
<td>Current</td>
<td>First choice source</td>
<td>BBC News (82%)</td>
</tr>
</tbody>
</table>
Table 3: Comparative statistics of specific web searching behaviours for English-based and Chinese-based data sets

<table>
<thead>
<tr>
<th>Multimedia Information</th>
<th>Main Reasons</th>
<th>Reliable (89%)</th>
<th>Fit reading habit (57%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First choice source</td>
<td>Youtube (63%)</td>
<td>Tudou or Youku (60%)</td>
<td></td>
</tr>
</tbody>
</table>

According to the results shown above, when searching for a specific kind of information, there are many differences between the two survey and also share a few common points. Firstly, English-based survey showed that respondents tend to choose Google and Wikipedia as their first choice of information source when search for factual information, but Chinese-based survey showed a different picture that Baidu was more popular among Chinese web users in this section due to their comprehensive service including Baidu encyclopaedia and natural language answer service called Zhidao; additionally, the main factors impact on users’ preferences were different as showed in the table 2. In terms of current information, the Portal websites news sections such as BBC and Sina are both the most popular among users in the two different languages. It should be noticed that the reasons for their choice was different, Chinese users tend to view “fit their reading habit” as a vital reason, while English-based survey showed “reliable” was the main reason for users to choose information sources among various kinds of services. In the end, the comparisons between the multimedia information showed that the video sharing websites where the contents shared by users has become the most popular forms of information source in this section, it also showed that Youtube among English-based users as well as Tudou or Youku among Chinese-based users are the most popular information sources for the same reason: the large collection of multimedia information.

5.2 Implications toward Web Search

As this study is aiming to validate whether the trend of web-based informational queries are changing, this study took a different strategy to ask people about their searching process. Apart from the end-up information sources showed above, the approaches to get to the end-up information sources listed above were also asked by
asking respondents to choose from five options which have been presented in the previous chapter: “a. Type URL into Address Bar”, “b. Type query into search engine”, “c. From my bookmark”, “d. From navigational websites”, “e. From personal homepage”. Among these options, the “b”, “d” and “e” options could be viewed as the navigational qui res generated from web search engines, since the navigational websites what respondents choose was literally owned by Baidu or Google.com and integrated in their “Homepage”. Besides, the reasons of causing this phenomenon were partly explained in this survey. Generally, the reasons of users’ choice of information sources in the different language Web environment were not the same, which also partly answered that why the most popular websites in English language were not as successful as their domestic Chinese counterparts because they know better about what factors impact on users’ preferences.

Consequently, the overall results showed that the web search engines were used more often to serve the navigational needs among Chinese language based Web environment, which proved the results concluded in the Tann’s study that a large group of queries have become partially navigational due to the emerging of the “super sites” in different domains on the Web; additionally, by the definition of “Transactional queries”, queries generated by users to search websites where information was not indexed by the web search engine and they could conduct a transaction could be divided into the categories of shopping information as well as social-related information, the results also showed that the transactional queries made by users more and more often recently.
Chapter 6 Conclusion

6.1 Conclusions

The objective of this research project was accomplished by conducting an investigation to validate a trend that the Web based informational queries are changing from a global perspective. More specifically, this trend refers to the Web based informational queries are partially shifting into navigational as well as transactional queries with a series of reasons in both English and Chinese Web environment.

In the beginning of this study, a comprehensive literature review was conducted in order to establish a foundation of informing theories and arguments related to this research project, with the theoretical basis made by Broder and further studied by Rose and Levinson as well as the latest work from Jansen, the taxonomy web search in terms of informational, navigational and transactional queries are well studied and discussed. Then the factors that impact on the web search engines especially the cultural and language factors were also discussed in order to address the objective of this study. In the end of literature review, the future trend of the web search engine was also paid attention to review from several aspects.

In the followed chapter, the methodologies adopted in this study were introduced as a mixture of quantitative research method and qualitative research method and the reasons to apply such strategy was also addressed in particular with the nature of research on web searching behaviours. Additionally, the data collection as a fundamental stage in this study was explained logically and smoothly stage by stage from the pilot stage, the design of questionnaire and the administration of survey.

According to the data collected after the distribution of the questionnaire, the certain findings were presented in the chapter four along with the discussions for each question. The main points could be summarised as following: Firstly, respondents’ experiences on web and search engine were sufficient in this research project; secondly, web search engines were widely used among the surveyed Chinese students but they tend to use more often with domestic brand web search engines such as Baidu for searching Chinese information on the web; thirdly, there are many
“super sites” are also existing in Chinese web users’ mind when they searching for information in various kinds of domains; finally, the informational queries of web search engine are shifting due to the existences of different “super sites” and becoming navigational queries as well as transactional queries with the reasons behind web users’ intentions.

In the chapter five, the trend of web based informational queries are changing partially into navigational as well as transactional queries in both English and Chinese language Web environment due to the emerging of the “super sites” in different domains on the Web was proved after the survey of the Chinese web users in this study and the comparisons between Tann’s study and this project.

After analysing the data collected from the Chinese students and then compared the findings with the previous English-based study conducted by Tann, this study not only confirmed the trend that Web-based informational queries are changing happened both in English language Web environment and Chinese language Web environment with the differences between cultural and language, it also explained the reasons of this trend were partly because of the emerging of the “super sites” in different domains on the Web since the new era of Web 2.0. This trend indicates a major challenge for the Web search engine industry in both English and Chinese Web environment, which web users’ intentions behind their information needs are continuing changing at present and the web search engine should not neglect people’s shifting needs. The possible solutions in light of this trend should be paid more attentions not only just introduce the personal Homepage and the navigational websites services, the more intelligent services also should be made from the nature of web search engine, which is the query typed into the web search engines from users to fulfil their “known intentions”.

6.2 Limitations and Recommendations

Although this study successfully reached the predefined objective, some flaws and gaps were found during the process of reviewing the project. Mainly, the problems were lied in the methodologies used in this research, there was limited population to participant in this study, which was partly due to the small scales of target group and
the time period was during the summer holiday; indeed, it also should be pointed out that by only asking respondents’ preferences of information sources and the reasons behind their choice, the answers were not as accurate as their behaviours in the real setting when they searching information on the Web in different domains. Subsequently, there were some minor inconsistencies in theory at findings, for instance, the relationship between cultural differences and the reasons of web users’ preferences of information sources was not discussed deeply and explained as clearly as expected due to the complexities of the nature of national culture and language; instead, the reasons of users’ preferences were compared between the two study to present the dynamic pictures of Web searching.

People are spending more time on the World Wide Web than ever at present and web search engines are also playing increasingly essential role in people’s daily life. Hence, it is imperative that the areas pertinent to future study should address how to improve the effectiveness of search engines for a varied set of tasks to better understand people’s intentions behind their changing information needs, some preliminary works have been starting recently such as the attempt to automatically classify Web-based queries on the basis of user intent by incorporating data mining categorisation techniques (Kathuria, A. et al., 2010). As a final point, it would be worthwhile to expand the study to more languages worldwide to further understand the users in different religions with more detailed Web searching experiences and then conduct a cross-cultural investigation to gain a more dynamic map of Web search industry.


Kathuria, A. et al. (2010). "Classifying the User Intent of Web Queries Using K-Means Clustering". Internet Research, **20** (5).


Appendix

Questionnaire

Information Sheet

I am a Chinese student currently take my master's degree in the Department of Information Studies.

This questionnaire is part of my final dissertation. The purpose of this questionnaire is to learn Chinese users' experiences of using search engine for accessing various kinds of Chinese language information.

The questionnaires is designed by Google Document and then distributed via university’s email system.

It will take your 10-15 minutes to finish this questionnaire.

I am very appreciated for your help. Hope you enjoy your summer holiday!

你好，非常抱歉占用您假期宝贵的休息时间。

我是谢菲尔德大学信息研究学院的一名中国学生，现正攻读硕士学位。这份问卷调查是我毕业论文的一部分，目的在于了解中国学生利用搜索引擎获取各类中文信息的基本经验，并由此与此前一份研究已知的外国学生获取英文信息方式进行对比与分析，能有助于我们更好的了解未来网络搜索趋势。

我保证这份问卷是完全的匿名问卷，并且您的任何个人信息都将得到保护。

在此诚恳希望同学能抽出 10-15 分钟时间完成这一份调查问卷，在此不胜感激您的帮助。
Part I. Initial Questions (基本问题)

Q1. What is your gender(请问您的性别是)?
A. Male 男性
B. Female 女性

Q2. What age group did you in (请问您处于哪个年龄段)?
A. 18-22
B. 23-25
C. 26-30
D. over 30

Q3. What course do you currently take (请问您正在攻读何阶段课程)?
A. Pre-Undergraduate Course 大学预科课程
B. Undergraduate Course 大学本科课程
C. Pre-Postgraduate Course 硕士预科课程
D. Postgraduate Course 硕士课程
E. PhD Course 博士课程

Q4. Please choose your time spend on the Internet per week (请选择您每周利用互联网的时间大概是多少)
A. Less than 10 hours 每周少于 10 小时
B. 10-20 hours 每周 10-20 小时
C. 20-30 hours 每周 20-30 小时
D. More than 30 hours 每周超过 30 hours

Q5. How often do you use search engine access Chinese language web information (关于您使用搜索引擎搜索中文信息的次数)?
A. Less than 5 times per week 每周少于 5 次
B. 1-3 times per day 每天 1-3 次
C. 4-6 time per day 每天 4-6 次
D. Frequently in everyday 每天经常使用
Part II. General searching behaviours questions (基本搜索习惯调查)

Q6: Please tick which kind of information did you mainly seeking on the web.
(请问您主要关注哪几类中文信息)
A. Current information (新闻时事)
B. Factual information (基本常识)
C. Academic Information (学业相关)
D. Multimedia Information (影音下载)
E. Shopping information (网上购物)
F. Social-related information (社交相关)

Q7: Which is the first choice when you search for Chinese language information?
(哪一个搜索引擎是您搜索中文信息的首选)
A. Baidu (百度)
B. Google.com
C. Sousou (搜搜)
D. Bing (必应)
E. Yahoo Yisou (雅虎中文)

Q8: Please tick the reason why choose the above search engine for Chinese information
(请选择您使用以上中文搜索引擎的原因)
A. Easy to use (方便易用)
B. More reliable (更为可靠)
C. More updated information (更新及时)
D. High-quality search result (搜索质量优秀)
E. Faster, just save the time (节省时间，搜索速度快)
F. No language barrier (没有语言障碍)
G. Good layout (界面亲切)
Q9. For Chinese language information websites, which kinds of source would you prefer?
(您更趋向于选择那一种中文信息网站来源)
A. Domestic Brand (本土网站品牌)
B. Foreign Brand (外来网站品牌)

Q10: Please tick the reasons about your choice
(请您勾选您以上选择的理由)
A. Easy to use (方便易用)
B. More reliable (信息更可靠)
C. Fit for my reading habits (符合我的阅读习惯)
D. Faster, just save the time (速度快，不需要长时间等待载入)
E. Good enough, no language barrier (够用就好)
F. Great interface (界面亲切)
Part III. Specific Questions (分类信息问卷)

Type 1: Chinese Factual information (中文基本常识信息)
When searching for Chinese factual information, e.g. historical data or documentary.
(此类信息包括如何获取基本常识信息，如历史资料、人物志或一些基本地理相关信息。)

Q11: Which is the first source used when searching for factual information?
(请选择您获取中文基本常识信息的首选信息来源)
A. Baidu (Baike, Zhidao) (百度(包括百度知道，百度百科))
B. Google.com ()
C. Wikipedia Chinese version (维基中文)
D. Chinese Portal Websites (e.g. Sina, Sohu, 163) (中文门户网站 (eg. 新浪, 搜狐,163))
E. Personal Blog (个人博客)

Q11(*) : Please specify your method to access above source
(请选择您到达以上所选项的主要途径)
A. Type URL into Address Bar (直接输入网址)
B. Type query into search engine to get there (通过搜索引擎再转到您的首选信息来源)
C. From my Bookmark (收藏夹直接访问)
D. From navigational websites (e.g.: hao123 or Very CD) (导航网站 (如 hao123, Very CD 导航等))
E. Personal Homepage (e.g.: My Baidu or iGoogle) (个性化主页 (如我的百度, iGoogle))

Q12: Please tick the reason why choose the above source
(对于您获取中文基本常识信息的首选来源，敬请选您的选择理由)
A. Accurate and in-depth enough (准确并有足够深度)
B. More convenient to locate information (方便定位)
C. I need high-quality information (我需要高质量的信息)
D. Fit for my reading habits (文章符合阅读习惯)
E. Like to participate in editing contents (能够参与编辑和讨论)
F. More updated information (信息更新及时)

G. Just save time (节省时间)

Q13. How do you evaluate the process of finding factual information?
(您是否对这类信息搜寻过程得心应手)
1. Very Difficult (无从下手)
2. Difficult
3. Normal
4. Easy
5. Very Easy (得心应手)

Type 2: Chinese Current information (中文时事新闻信息)
When searching for Chinese Current information, e.g. World cup news or domestic news.
(举例说明：此类信息包括获取当前即时新闻报道，如本地新闻、世界杯新闻或世界各地新闻报道。)

Q14. Which is the first source used in regards to current information?
(请选择您搜寻中文时事新闻信息的首选信息来源)
A. Baidu News (百度新闻)
B. Google News Chinese version (Google 中文新闻)
C. Sina News (or 163, ifeng) (新浪新闻或其他中文门户网)
D. RSS News (RSS 新闻)
E. Social Websites (e.g. QQ or Renren) (社交网站（如人人、QQ 新闻）)

Q14 (*): Please specify your method to access above source
(请选择您到达以上所选项的主要途径)
A. Type query into search engine to get there (通过搜索引擎再转到您的首选信息来源)
B. From my Bookmark (收藏夹直接访问)
C. From navigational websites (e.g.: hao123 or Very CD) (导航网站（如 hao123, Very CD 导航等）)
D. Personal Homepage (e.g.: My Baidu or iGoogle) (个性化主页（如我的百度，iGoogle）)
Q15: Please tick the reason why choose the above source

A. Always find necessary information (总能查找到相关必要信息)
B. Fit for my reading habits (更方便浏览阅读)
C. Good layout and content (内容合理易懂)
D. More reliable (准确可信)
E. I can easily make comment (乐于参与评论)
F. Save time (节省时间)
G. In-depth enough (深度报道，更多分析)
Other

Q16. How do you evaluate the searching process of finding the current information experience?

1 Very Difficult (无从下手)
2 Difficult
3 Normal
4 Easy
5 Very Easy (得心应手)

Type 3: Chinese Shopping information (中文购物信息)
When searching for shopping information, e.g.: digital camera or books.

Q17. Which is the first source used in regards to shopping information?

A. Taobao (淘宝网)
B. Dangdang (当当网)
C. BBS (各类论坛)
D. Baidu (百度)
E. Google.com

Q17 (*): Please specify your method to access above source
(请选择您到达以上所选项的主要途径)
A. Type query into search engine to get there (通过搜索引擎再转到您的首选信息来源)
B. From my Bookmark (收藏夹直接访问)
C. From navigational websites (e.g.: hao123 or Very CD) (导航网站（如 hao123, Very CD 导航等）)
D. Personal Homepage (e.g.: My Baidu or iGoogle) (个性化主页（如我的百度，iGoogle）)
E. Type URL into Address Bar (直接输入网址)

Q18. Please tick the reason why choose the above source
(对于您获取中文购物信息的首选来源，请勾选您的选择理由)
A. Convenient to purchase (一站式服务，方便购买)
B. Comparable price information (准确定位商品，透明且可比较的价格)
C. Just save time (节省时间)
D. Great shopping experience (良好的购物体验，界面亲切)
E. More reliable, I like to get more information on people's views (感觉更可靠，人们的评论很重要)
F. Fit for my shopping habits (符合我的购买习惯)

Q19. How do you evaluate the searching process of finding the shopping information experience?
(您是否对此类信息搜寻过程得心应手)
1 Very Difficult (无从下手)
2 Difficult
3 Normal
4 Easy
5 Very Easy (得心应手)
Q20. Please enter your suggestions to address any issues about above questions?

(请填写您对以上问卷的意见和建议)

Type 4: Multimedia information (影音娱乐相关信息)
When searching for multimedia information, e.g.: audio or video.
(举例说明：此类信息包括获取音频、视频相关信息。例如视频短片，电影、歌曲等。)

Q21. Which is the first source used in regards to multimedia information?

(请问您获取影音信息的首选信息来源是)
A. Multimedia Share websites (e.g.: Tudou, Youku) (视频分享网站（如土豆，优酷）)
B. Social sharing websites (社交分享网站（如人人，开心，QQ空间）)
D. Portal Websites (e.g. Sina, 163) (中文门户网站（如新浪，163等）)
E. Baidu (百度)
F. Google.com

Q21(*): Please specify your method to access above source

(请选择您到达以上所选项的主要途径)
A. Type query into search engine to get there (通过搜索引擎再转到您的首选信息来源)
B. From my Bookmark (收藏夹直接访问)
C. From navigational websites (e.g.: hao123 or Very CD) (导航网站（如hao123, Very CD导航等）)
D. Personal Homepage (e.g.: My Baidu or iGoogle) (个性化主页（如我的百度，iGoogle）)
E. Type URL into Address Bar (直接输入网址)
Other

Q22. Please tick the reason why choose the above source

(对于您获取影音信息的首选来源，请勾选您的选择理由)
a. Enough collection of multimedia information (收藏丰富齐全)
b. Easy to locate target information (准确定位到目标影音信息)
c. Fast enough (速度够快)
d. Good image quality (画质或音质较好)
e. Like to discuss with peers (能够与他人进行交流讨论)
f. Update quickly (更新及时，内容新颖)

Other

Q23. How do you evaluate the searching process of finding the multimedia information experience?
(您是否对此类信息搜寻过程得心应手)

1 Very Difficult (无从下手)
2 Difficult
3 Normal
4 Easy
5 Very Easy (得心应手)

Type 5: Social-related information (社交类相关信息)
When searching for social-related information.
(此类信息包括获取社交、兴趣爱好等相关信息。)

Q24. Which is the first choice you search for social related information?
(请问您获取社交相关信息首选信息来源是)
A. Renren (Xiaonei) (人人网（校内）)
B. Kaixin (开心网)
C. Q-Zone
D. Personal Blog (各类博客)
E. Douban (豆瓣网)
F. Microblog (eg: Sina Microblog) (微博（如新浪微博等）)
G. Baidu (百度)
H. Google.com

Q24 (*): Please specify your method to access above source
(请选择您到达以上所选项的主要途径)
A. Type query into search engine to get there (通过搜索引擎再转到您的首选信息来源)
B. From my Bookmark (收藏夹直接访问)
C. From navigational websites (e.g.: hao123 or Very CD) (导航网站（如 hao123, Very CD 导航等）)

D. Personal Homepage (e.g.: My Baidu or iGoogle) (个性化主页（如我的百度，iGoogle）)

E. Type URL into Address Bar (直接输入网址)

Q25. Please tick the reason why choose the above source
(对于您获取社交相关信息的首选来源，请勾选您的选择理由)

a. Easy to connect with friend (方便建立联系)

b. Real and reliable (真实可信)

c. Fast and updated quickly (载入速度快，更新及时)

d. I like their Interesting applications (丰富有趣的各类应用)

e. Integrated instant message function (融合即时信息交流平台)

f. Great interface (界面亲切)

Other

Q26. How do you evaluate the searching process of finding the current information experience?
(您是否对此类信息搜寻过程得心应手)

1 Very Difficult (无从下手)

2 Difficult

3 Normal

4 Easy

5 Very Easy (得心应手)

Part IV. Please enter your suggestions for Chinese language websites to improve their service.
(如果您有任何的意见和建议，请在下面方框内自由发表，您的意见和建议将会被虚心采纳。再次表示谢意.)

Please write down here.