A Usability Approach to Improving the User Experience in Web Directories

Jiang, Nan

ISSN 2043-0167

A Usability Approach to Improving the User Experience in Web Directories

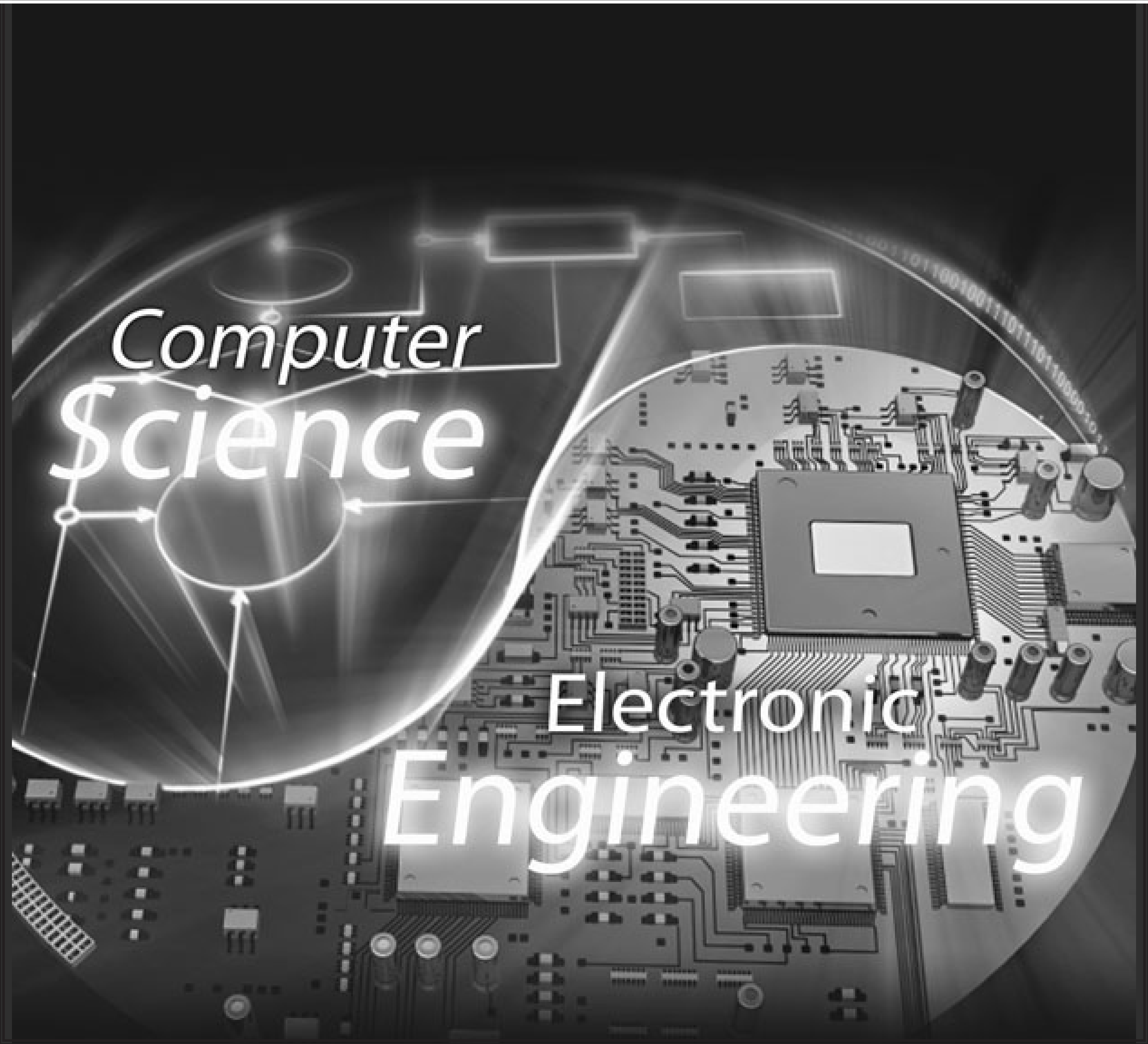
Nan Jiang

Queen Mary
University of London

EECSRR-09-02

December 2009

School of Electronic Engineering
and Computer Science



A Usability Approach to Improving the User Experience in Web Directories

Nan Jiang

Submitted for the degree of Doctor of Philosophy

Queen Mary, University of London

2009

Abstract

Web directories are hierarchically organised website collections that offer users subject-based access to the Web. They played a significant part in navigating the Web in the past but their role has been weakened in recent years due to their cumbersome expanding collections. This thesis presents a unified framework combining the advantages of personalisation and redefined directory search for improving the usability of Web directories.

The thesis begins with an examination of classification schemes that identifies the rigidity of hierarchical classifications and their suitability for Web directories in contrast to faceted classifications. This leads on to an Ontological Sketch Modelling (OSM) case study which identifies the misfits affecting user navigation in Web directories from known rigidity issues. The thesis continues with a review of personalisation techniques and a discussion of the user search model of Web directories following the suggested directions of improvement from the case study. A proposed user-centred framework to improve the usability of Web directories which consists of an individual content-based personalisation model and a redefined search model is then implemented as D-Persona and D-Search respectively. The remainder of the thesis is concerned with a usability test of D-Persona and D-Search aimed at discovering the efficiency, effectiveness and user satisfaction of the solution. This involves an experimental design, test results and discussions for the comparative user study.

This thesis extracts a formal definition of the rigidity of hierarchies from their characteristics and justifies why hierarchies are still better suited than facets in organising Web directories. Second, it identifies misfits causing poor usability in Web directories based on the discovered rigidity of hierarchies. Third, it proposes a solution to tackle the misfits and improve the usability of Web directories which has been experimentally proved to be successful.

Acknowledgements

This thesis is dedicated to Jin Wang, my mother who showed her love in both emotionally and financially ways. I would also like to thank to my two aunts, Qin Wang and Jie Wang, for their help during my stay in London.

I am very grateful for the encouragement and advice from my two supervisors, Jon Rowson at Queen Mary and Jane Reid at Queen Mary throughout the whole research period. Especially, for Jon's overall support and Jane's advice on conducting experimental studies. I would also like to thank Kathy Abbott, our subject librarian for sharing her expertise in cataloguing and classification. Thanks are also due to the time and interest of all students who participated the experimental study. Without all of you, this work would have not been possible.

Thanks to the online communities including the beginners mailing list for Perl and MSDN: Microsoft Developer Network for sharing advice and thoughts in the implementation.

Lastly, to Jingyu Bi, my wife for the understanding and patience she gave to me.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Directory

→ ***noun*** (pl. **directories**)

1. a book listing individuals or organizations alphabetically or thematically with details such as names, addresses, and telephone numbers.

• a board in an organization or large store listing names and locations of departments, individuals, etc. • (Computing) a file which consists solely of a set of other files (which may themselves be directories).

- ORIGIN late Middle English (in the general sense ‘something that directs’): from late Latin directorium, from director ‘governor’, from dirigere ‘to guide’.

[*Oxford Dictionary of English (Revised Edition) 2005*]

1.1 The General Definition of Web Directories

Web directories are a variety of directories on the World Wide Web which are only varied from traditional ones in the form of data collection and management. In details, Web directories contain specific types of online resources[1] with titles and descriptions and use an interactive manual inclusion mode for gathering them. For example, users can suggest or submit resources to specific categories while editors reserve the right for approvals. These differ from traditional directories which contain only contact details for individuals and businesses and use editors for maintenance and have been generally accepted for defining Web directories. For instance, Wikipedia (2008) states a Web directory's relationship with traditional directories as “*a directory on the World Wide Web*” and identifies its speciality in “*linking to other web sites and categorizing those links*”. It also points that a Web directory “*often allows site owners to submit their site*

1 Such types of resources include a whole website address (e.g., http://www.bbc.co.uk/ in “News and Media” of DMOZ) , a section or a micro site of a website (e.g., http://www.bbc.co.uk/weather in “Weather (UK) of DMOZ”) and even a web page presenting a specific topic (e.g., http://www.bbc.co.uk/bbcfour/audiointerviews/profilepages/lecorbusierc1.shtml in “Le Corbusier, Charles-Edouard (Architects, Arts)” of DMOZ).

for inclusion, and have human editors review submissions for fitness". Similar views can be also found in leading Web directories. For example, Yahoo! Directory (2008) describes itself as "*a human-created and maintained library of web sites organized into categories and subcategories*". About.com "*organizes Web sites by subject, and is usually maintained by humans instead of software*" (Boswell, 2004). Open Directory (2008) says it is a "*human edited directory of the Web and the purpose of the ODP is to list and categorize web site*".

Thus, we could generally define a Web directory as
"*a human compiled and maintained, subject-based resource collection which contains classified and reputable websites in hierarchically aligned categories with cross references and aims to guide Internet user to navigate through the Web.*"

Consider the speciality of Web directories, there are two kinds of them: *general directories* and *specialised directories*. General Web directories, also known as generic directories, try to organise different kinds of online resources as much as possible so as to give users a subject-based access of the whole Web while specialised directories only focus on specific aspects of the Web. For example, The Open Directory Project[2] is a comprehensive Web directory classifying the Web into sixteen general domains whereas Chef Moz[3] is a specific directory for gathering online information for restaurants in the world. In this thesis, we focus on general Web directories and therefore use "Web directory" and "Web directories" to refer generic Web directories.

1.2 The Past and Present of Web Directories

1.2.1 The Rise of Web Directories (1994 – 1998)

2 http://dmoz.org/
3 http://chefmoz.org/

Web directories were the most popular navigation service in the period of "the growth of WWW" (1992 - 1995). The WWW Virtual Library[4] (2008), which is the oldest online catalogue launched by Tim Berners-Lee, is even two years older than CERN's public release of the World Wide Web in 1993. Soon, the first commercial Web directory Yahoo![5] (the acronym of *Yet Another Hierarchical Officious Oracle*) was launched by Jerry Yang and David Filo in 1994 (Figure 1.1). Later, Best of the Web[6] and Lycos' Top 5% (1994 –2000), two other favourite Web directories in that time, joined the group. These directories, more or less, came for the same intention, which was to become a high quality Web guide consisting of only favourable and reputable websites to help people navigate on the Web.

Web directories gained a steadily growth in the commercialisation of WWW between 1996 and 1998. Many current big names were born in this period. For example, Starting Point Directory[7] and LookSmart[8] were launched in 1995 and later Open Directory Project in 1998.

4 http://vlib.org/
5 http://dir.yahoo.com/.
6 http://botw.org/
7 http://stpt.com/
8 http://www.looksmart.com/

Figure 1.1 Yahoo! Directory in The Year of 1994

Retrieved August 19, 2008 from http://web.bilkent.edu.tr/History/yahoo/

It is not difficult to speculate why Web Directories gained their popularity quickly. First, Web directories were built as a portal to guide users navigate on the Web. To explore "a new world"(Shneiderman et al.1998), this was the most demanding tool when there was the lack of adequate domain knowledge. Second, the Web had a relatively small size and low diversity in terms of subjects at that time so that building directories was an effective and easy approach. Figures in RFC 2235 (Zakon, 1997) showed that there were only 2,738 websites when Yahoo! was born where most of them were governmental websites (Zakon, 2006). Although the Web inflated rapidly after 1996, websites still came from relatively constraint sectors due to the private ownership of domains. Last, search engines (e.g., Wandex and Aliweb founded in 1993) in that time, compared to these relatively well-constructed Web directories, were somewhat primitive[9] to use. Clues can be found through the "dependency" of those search engine forerunners like Lycos (1994), Excite (as Architext in 1994), Infoseek (1994) and AltaVista (1995) when they announced search engines and their own Web directories together. For example, "Lycos' Top 5%" and "Sites by Subject" from Lycos, "Explore Excite" from Excite and "Explore these popular Infoseek Select topics" from Infoseek. This also consequentially made the trend of using "Web portal", "Web directory" and "search engine" interchangeably in those times.

The Web expanded enormously and received strong recognitions as a result of its commercialisation movement since 1996 (Figure 1.2). By the end of 2001, the total number of websites had reached 36,276,252, over ten thousand times more than the period of mid 1994 (Zakon, 2006).

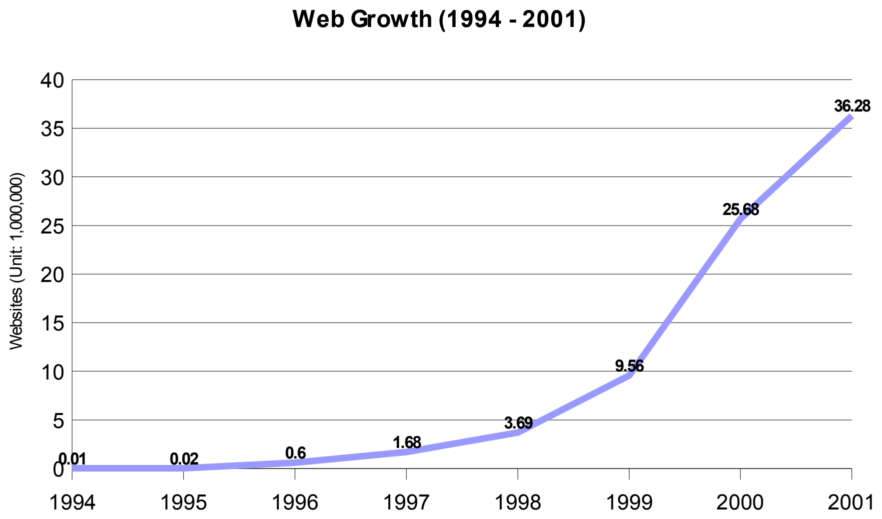


Figure 1.2 Web Growth (1994 – 2001)

Retrieved August 19, 2008 from Hobbes' Internet Timeline

Criticism of Web directories firstly came out from an annual submission survey of Yahoo! (Sullivan, 1997a, 1997b, 1997c) where the acceptance rate of website inclusion dropped significantly year by year while the demand of inclusion was increasing (Table 1.1).

Are You Listed?	Responses	No	Yes
Submitted in 1995	5	0%	100%
Submitted in 1996	25	60%	40%

9 Earlier search engines using primitive Web crawling and spidering techniques, were only able to index a limit number of static webpages. Such search engines' performance also often relied heavily on training samples of indexing.

Submitted in 1997	132	77%	23%

Table 1.1 A Summary of Yahoo! Annual Submission Survey

Retrived on August 19, 2008 from http://searchenginewatch.com/showPage.html?page=2165571

Similar issues were also found in other directories as they could not scale to the rapidly growing Web. For example, The Open Directory collected only 493,711 websites by January 1999 when the volume of the Web reached 4,062,280 websites in the same period[10]. Nevertheless, the inclusion issue seemed like no big deal to normal Web users as Web directories were still reputable navigation services. According to Yahoo! (1997a, 1997b, 1998), the largest commercial Web directory ranked No.1 website for navigation and attracted the most audience during that time. Additional evidence could be found when Sullivan (1998) claimed that it was the centre stage of Web directories which helped LookSmart gain high traffic growth after launching a directory service called "Relevant Knowledge's Top 25 web sites". The tendency also made Lycos transform itself into a vertical Web directory from a search engine on April, 1999 (Sullivan, 1999; Kerber, 1999).

1.2.2 The Fall of Web Directories (1999 – 2006)

The "dot-com bubble" (Wikipedia, 2008) helped the Web expand to host 36,276,252 websites by the end of 2001 which is almost 10 times that of 1998. However, Web directories did not respond this growth very well. Using DMOZ as an example, its website coverage declined to 8.3% of 2004 from 27% of 1998. In order to deal with the increasing complaints about the inclusion delay, pay inclusions allowing editors to treat commercial sites with priority, was introduced by Yahoo! in 1999 (Kane, 1999). The mechanism soon became a popular solution for commercial directories but its drawback was almost as significant as its benefit: editors lost the opportunity of collecting "reputable" websites as usual because they had to give their priority to those paid ones. To some extent, pay inclusions pushed Web directories to the side of yellow pages from the side of quality Web guides. Moreover, with more and more websites and categories

10 Source data are gathered from http://www.archieve.org and Hobbes' Internet Timeline.

being added and created into Web directories, normal Web users were unsatisfied with the poor accessibility in Web directories' labyrinth-like categories. On the other hand, search engines were improved significantly in terms of the size of index and search quality. For example, Google claimed it had indexed over 1 billion webpages in October, 2000 (data source: Internet Archive). This not only increased the popularity of search engines but also made Web users choose search engines as their portals for locating websites. Forrester Research (2000) reported that " *in UK, search engines are the leading way users locate web sites*" and "*Search engines are the top way consumers find new web sites online, used by 73.4% of those surveyed*" (VanBoskirk & Li, 2001). Eventually users seemed to forget about the other methods of obtaining information on the Web. Evidence for this can be found in the statistical data provided by Web monitoring company Alexa Internet. In websites offering both search engine and directory services, directories received only 3.6% of user attentions compared to 96.56% for search engines (Table 1.2).

Websites	**Reach* of directory part**	**Reach* of search engine part**	**Ratio (Directory/Search Engine)**
Open Directory	232	0 (n/a)	n/a
Yahoo!	2,854	28,540	10%
Google	2,671	189,641	1.4%
MSN	2,454	9,814	25%
Lycos	47	538	8.7%
AltaVista	0 (Yahoo alliance)	3,621	0%
Look Smart	14.15	156	9.1%
Total	**8,272.15**	**232,310**	**3.6%**

* *Reach stands for 3 months average visits per million users. Source: Alexa data of 2006 Q3*

Table 1.2 A Comparison of the Average Visits for Major Web Directories

1.2.3 The Need for Web Directories (2006 -)

Web directories are now commonly used for SEO[11] (search engine optimisation) (Hawley, 2005) by professional users (e.g., website masters and digital agencies) rather than for guiding normal users to new websites. Does the role change indicate that Web

11 The idea behind it is, links from reputable sources like Web directories will improve the ranking of websites in major search engines.

directories are not important to normal users any more as they have search engines? This is unlikely to be true.

The main reason is that Web directories and search engines utilise different information seeking behaviours, *browsing* and *searching* respectively. Browsing is a semi-directed or semi-structured behaviour (Ellis, 1989; Ellis et al., 1993; Ellis & Haugan, 1997) which assists user to investigate information with some vague ideas or even with a "blank mind" and then extend their personal interests gradually (Kuhlthau, 1991, 1993, 1994, 1997). Palay & Fox (1981) point that this can be only achieved within a well structured information source where all information is clearly organised and classified. Since the Web is not formed in such a way, Web directories were introduced as the only available collections of the Web so as to bring the possibility of browsing. For example, a user can browse the top-level categories of Yahoo! Directory and choose one of them to specify his interests level by level until he finds desired a website collection. Searching, in another way, is an intuitive behaviour which requires user to initialise the process with specific purposes and search engines were introduced for supporting this kind of process. For example, if a user is interested in some concrete facts on the Web, he can describe them with a set of keywords and search them into Google. Note he can only get satisfying results when he has specific interests and clear goals. In addition, the different nature presented by browsing and searching also determines that Web directories and search engines target on different search needs in terms of the nature of results. That is, Web directories are ideal for finding a particular set of resources (as a specific category) and search engines are for a specific piece of information (as a specific webpage). According to these, Web directories are still a unique service which are unlikely to be replaced by search engines.

In details, Web directories have some key characteristics where search engines do not have.

- ***Subject-based access***

  Web directories have been acknowledged for providing subject-based access to the Web through their clear and well-structured information representations which not only allow user to start with vague ideas but also help them specify their interests in logical orders and relationships. A good example is given by Google directory (Google, 2008): if a user wants to find out about a US football club called "Lion", he could go to "United States" then click "Sports" and "Football Clubs".

- ***Human understandable organisations***
  Web directories organise resources by classifying them into hierarchically arranged categories based on the likeness and distinctness of subjects which are semantic and highly adapted to human understanding.
- ***Quality website collections***
  Web directories offer quality and reliable website collections where only reputable resources get reviewed and collected as a result of their strict submission guidelines (Wikipedia, 2008; Sherman, 2000).

1.3 Research Question

Web directories played a significant part in helping users navigate the Web in the past but their role has been weakened in recent years. The rise and fall of Web directories suggests that the number of resources collected or indexed could be important for the navigational experience when the Web keeps expanding enormously. However, unlike the mechanism used by search engines, Web directories only need to be representative rather than comprehensive. Moreover, unlike search engines which have invisible structures, the level of communication and interaction with visible structures offered by Web directories seems to be a definite factor to affect the user browsing and favourite. This becomes even more serious a problem when the number and depth of categories increases even more. The user experience of accessing the Open Directory containing over 590,000 categories to obtain useful information in 2008 is hugely different compared to accessing the same directory with only 77,773 categories in 1999 (data source: http://web.archive.org/). As no relevant research has been done in considering the effect of interaction on the user experience in Web directories, we are therefore driven to performing a usability study for investigating and understanding the decline of Web directories in this direction. Thus, our research question is,

Is there a particular usability solution that can be used for improving user experience in navigating Web directories so as to make Web directories more useful?

1.4 The Remaining Chapters

The main research work is presented in the next six chapters.

The background study is carried out in Chapter 2, which aims to reveal the link between the rigidity of hierarchical classification schemes and the representation of Web directories in terms of user navigation. Two major classification schemes and their suitability for Web directories are also discussed in this chapter.

A case study which applied Ontology Sketch Modelling (OSM) to conduct a usability inspection, is reported in Chapter 3 and its results are discussed for discovering the misfits between Web directories, their representations and their users. Two possible directions for improving these misfits are then introduced.

A unified framework consisting of an individual personalisation model and a redefined search model is proposed in Chapter 4. Highlights of its implementation, D-Search and D-Persona are also explained in this chapter.

The experimental design of D-Search and D-Persona is introduced in Chapter 5 where a comparative user study comparing the performance of Open Directory, Google and D-Search and an open user study for D-Persona are explained in details.

Results and discussions of the comparative study are reported in Chapter 6 and the final conclusion and future directions of the research are drawn in Chapter 7.

Chapter 2 Background Study

"Classification scheme is the descriptive information for an arrangement or division of objects into groups based on characteristics which the objects have in common."
(ISO/IEC 11179-2)

2.1 Introduction

Web directories classify information resources on the Web with pre-established classification schemes. In the above definition given by ISO/IEC, "an arrangement" refers to the organisational structure of Web directories, "objects" mean URLs and "characteristics ... in common" reflect the classificatory view of the likeness or sameness of these URLs. This interpretation could help us understand how Web directories organise and represent Web knowledge in order to guarantee successful user navigation.

Generic Web directories like Yahoo! Directory, Open Directory or Best of the Web use hierarchical classification schemes which feature high matureness, good overall acceptance and wide economic applicability (Koch, 1997). Suppose you are looking for art museums in London in the Open Directory. First, you need to choose one out of sixteen top level categories to start your browsing. Then you have to narrow the selected subject down to one of eleven secondary categories (supposing you chose "Regional" in the first step) and then another sub category again until you find the right one at the 7th level, which is "Regional: Europe: United Kingdom: England: London: Arts and Entertainment: Museums". Since a subject (class) in the hierarchical classification is organised in a fixed broad-to-narrow relationship, users have to make a series of trial-and-error attempts to find out this relationship so as to retrieve the subject. That is, when browsing in a hierarchical classification, whenever a user has made some wrong choice on some level, they have to go back to this level and make another choice. If this is still not working, they have to go back to even higher levels. With the never-ending expansion of categories in large Web directories, the trial-and-error attempts would become more difficult and also cause the complexity of navigation to increase dramatically. The Open Directory contains 4,612,597 sites in over 590,000 categories at the time of writing[12] compared to 493,711 sites in 77,773 categories on the 7th of May, 1999[13]. There is clearly a huge navigational difference in accessing the same subject in this directory between now and past.

Studies in LIS (Library and Information Science) have found that such navigational difficulties, which are also called rigidity, were very common in large hierarchies where some predefined establishing principles to classify dynamically growing knowledge domains are used. In the book of 'Philosophy of Library Classification' (1951), Ranganathan, the founder of the Colon Classification described the rigidity in the following context:

"An enumerative scheme with a superficial foundation can be suitable and even economical for a closed system of knowledge......What distinguishes the universe of current knowledge is that it is a dynamical continuum. It is ever growing; new branches may stem from any of its infinity of points at any time; they are unknowable at present. They cannot therefore be enumerated here and now; nor can they be anticipated, their filiations can be determined only after they appear." (Ranganathan, 1951).

Ranganathan and others (Ranganathan, 1951; Wynar & Taylor, 1992; Taylor, 2000; Drengson, 1996; Kim et al., 2006) then suggested using faceted classifications to overcome this problem. Recent researches (Ellis & Vasconcelos, 2000; Bates, 2002; Broughton, 2002; Zins, 2002; Kim et al., 2006) also show that faceted classifications are efficient in organising dynamic online resources for their flexibility in accommodating new knowledge, supporting multidimensional views and presenting multiple

12 Data is collected from the homepage of Open Directory at http://www.dmoz.org on October 13, 2008.
13 Data is retrieve from Web archive on Oct 6, 2008 at http://web.archive.org/web/19990508064303/207.200.73.135/.

relationships. After some successful applications in organising e-commerce websites, the interest of using faceted classifications to organise Web resources has grown quickly (Denton, 2003; Adkisson, 2003; Lin, 2006; Kim et al., 2006). Thus,

Can we apply faceted classification schemes to Web directories so as to reduce the navigational difficulties and complexities in these hierarchically organised directories?

The question will eventually be answered in Section 2.3 through a comparison between hierarchically classification schemes and faceted classification schemes after both schemes have been examined in Section 2.2 and the current rigidity issue in hierarchical classifications is defined.

2.2 A Review of Universal Classification Schemes

Universal classification schemes that focus on organising the whole universe of knowledge are previously called library classification systems which were created for the overall needs of organising the massive book collections obtained in the national libraries[14]. Library classification systems generally play two roles. First, they facilitate subject access by allowing user to find out what works or documents the library has on a certain subject. Second, they provide a known location for the information source to be shelved and subsequently found. These two roles have become the basic requirement of all classification systems since then. There are two types of classification schemes, hierarchical classifications and faceted classifications. *Hierarchical classifications*, as seen in Web directories, divide subjects hierarchically from the most general to the most specific with enumerated classes and relationships. *Faceted classifications*, as seen in many e-commerce websites, organise subjects with mutually exclusive orthogonal facets (properties of objects).

Famous classification schemes include the Dewey Decimal Classification (DDC), the

14 Legal deposit is a legal requirement that one copy of every book published has to be submitted to a repository, usually a national library.

Universal Decimal Classification (UDC), the Library Congress of Classification (LCC), the Colon Classification (CC) and the Bliss Classification (BC and BC2). These schemes, especially the DDC and LCC, have deeply influenced the development of Web directory classification schemes (Steinberg, 1996; Vizine-Goetz, 1996; Koch, 1997).

2.2.1 Hierarchical Classification Schemes

Hierarchy, in Greek, ***Ἱεραρχία*** *(Hierarchia), derived from* ***ἱερός*** *– hieros, "sacred" and* ***ἄρχω*** *– archein, "rule, command".*

2.2.1.1 Definition

Pseudo-Dionysius Areopagita (the 5th century) is believed to be the first person to use this word in his works, "The Celestial Hierarchy" and "The Ecclesiastical Hierarchy" to denote the ruling powers in the Church. In fact, the history of this ordering concept is much older than the word itself. Our understanding of hierarchical classifications is inherited from Aristotle (384 BC – 322 BC) (Ackrill, 1963; Kwasnik, 1999), who posited that a unified whole comprised by all nature could be subdivided into classes, and each class further divided into subclasses. He stated that a natural dividing place of any given class is determined by the necessary and sufficient attributes for membership in the class and the classification process must follow an orderly and systematic set of rules of association and distinction. This general idea is alive in spirit in today's classification systems.

In general, a hierarchy is a system of entities partially ordered by some kind of *relations* *of subordination*. That is, each entity (except for the top entities) of the system is subordinate to a single other entity in some way. A hierarchy can link entities either directly or indirectly, and either vertically or horizontally. The only direct links in a hierarchy are to one's immediate superior or to one of one's subordinates. Indirect hierarchical links can extend "vertically" upwards or downwards via multiple links in the same direction. All parts of the hierarchy which are not vertically linked to one another can nevertheless be "horizontally" linked by travelling up the hierarchy to find a common direct or indirect superior, and then down again. There are two common relations of subordination, *type hierarchy* which presents a *type-subtype* relation (e.g., furniture: wardrobe; table; sofa; bed...) and *part hierarchy* which presents a *whole-part* relation (e.g., desktop computer: components; display; peripherals...). A hierarchical classification system for knowledge organisation usually adopts a *broader-narrower* hierarchy, which can be considered as a mixture of *type*- and *part*-hierarchies (Hjørland, 2008).

2.2.1.2 Classic Hierarchical Classification Schemes

The Dewey Decimal Classification (DDC), initially devised by Mevil Dewey in 1876 (Dewey, 1979), is now being used in 200,000 libraries worldwide as the world's most widely used library classification system. It is updated on an ongoing basis for keeping pace with knowledge and the current version includes DDC 22 in 2003 and Abridge Edition 14 in 2004 (OCLC, 2008). The DDC organises knowledge by disciplines or fields of study. At the broadest level, the DDC is divided into ten *main classes* with decimal notations from [000] to [900], which together attempts to cover the entire world of knowledge. Each main class is further divided into ten *divisions* and each division into ten *sections*. Hierarchies of the system are expressed through structure and notation. The *structural hierarchy* is also called "hierarchical force". That is, all topics (aside from the ten main classes) are part of all the broader topics above them so a subordinate class should only present a narrowed subject of its superordinate class. In the following example where the bold class indicate that it is a main class of the DOC and the decimal numbers in front of each class is its notation, "Photographs" is a subject of “Photography & Photographs” and both “Photographs” and “Photography & Photographs” belong to “Arts & recreation”. This broad-to-narrow relationship is also presented by the *notational hierarchy*, which is used for optimising the categorising and shelving process in libraries but has not seen in Web directories.

700 Arts & recreation
770 Photography & Photographs
771 Techniques, equipment, materials
779 Photographs

The main advantage of the DDC over other library classification rivals is its simplicity, attributed to the use of pure hierarchical decimal notation and a mnemonics system, which make it generally easy to use for most users.

The Library of Congress Classification (LCC) is another remarkable classification system which was influenced by the DDC and developed by Herbert Putnam with the advice of Charles Ammi Cutter (the founder of Cutter Expansive Classification) in 1897. It was specifically designed for managing the books submitted to the Library of Congress of the United States but is now being used by most research and academic libraries in the US and several other countries. The LCC is based on 21 main *classes* and each of them is arbitrarily assigned one of the letters A through Z excepting I, O, W, X and Y. Each main class is independently divided into various numbers of *subclasses* by individual experts in each area according to the demand of cataloging. The subclasses are assigned with one or two additional letters to the main letter and a set of numbers based on Cutter Expansive Classification (Cutter, 1962). In the following example, note that the bold class is a main class of the LCC, “Cameras” and “Photographic Processing” are the-same-level sub-subjects of “Photography” in “Technology”.

T	**Technology**
TR 1-1050	**Photography**
TR 250-265	**Cameras**
TR 287-500	**Photographic Processing**

The LCC is very technically oriented which could be seen from the above example where “photography” is described as a subject of technology rather than a general work of art as in the DDC. This orientation makes the LCC difficult for non-expert users. Other features of the LCC include country based subdivisions and extensive use of cross-referencing such as NT (Narrower Term), BT (Broader Term), RT (Related Term), SA (See Also), UF (Used For) and USE.

2.2.1.3 The Hierarchical Classification Schemes of Web Directories

Except for a few Web directories claiming their origins in traditional classification schemes (such as CyberDewey of the DDC (Mundie, 1995) and the WWW Virtual Library of the LCC), others tend to develop their homegrown schemes by combining the characteristics of the DDC and LCC (Koch, 2004). Consider two dominating Web directories, Yahoo! Directory and the Open Directory. Yahoo! Directory is similar to the DDC by employing 20 editors to develop its scheme constantly (Steinberg 1996) and the Open Directory is like the LCC by assigning 80,757 volunteering editors[15] for managing different categories from the point of classification consistency. Their likeness of the DDC and LCC are reversed in terms of user groups where Yahoo! Directory is targeted for commercial use like the LCC and the Open Directory, aims for general use as the DDC. Both directories use detailed subject division in each level of the classification which is similar to the LCC but they apply simple cross-referencing strategy similar to the DDC by only using “see also” and “@”. Therefore, we suggest to consider the DDC and LCC as a whole in order to understand the characteristics of hierarchical schemes.

2.2.1.4 The Essence of A Hierarchy

Regardless of their own characteristics, both the DDC and the LCC present a unique but strict order passed down from top to bottom, which can be only divided level by level but cannot be duplicated or overlapped at any level. This can be illustrated by using the

15 The figurate is based on the homepage of the Open Directory on August 19, 2008.

following example (Figure 2.1).

- **Publications**
    - **Magazines**
        - **Business Magazines**
            - **Weekly Magazines**
            - **Monthly Magazines**
        - **Consumer Magazines**
    - **Newspapers**
        - **Local Newspapers**
        - **National Newspapers**
        - **International Newspapers**
    - **Books**
        - **Comic Books**
        - **Textbooks**
    - **Research Journals**

Figure 2.1 A Sample Hierarchical Classification for Publications

Class inclusiveness

Inclusiveness is a property held by all classes containing subclasses in a hierarchical classification. The top-level classes in the classification are the most inclusive (broadest) classes for describing the knowledge domain and other classes are less and less inclusive (narrower) with the deeper and deeper levels. For example, 'publications' include 'magazines', 'newspapers', 'books' and 'research journals' and all their subsets while newspapers only include 'local newspapers', 'national newspapers' and 'international newspapers'.

Inheritance and transitivity

Class inclusion ensures that all attributes or properties of a given class in a hierarchy will be inherited by its subclasses and sub-subclasses. Whatever is true of 'publications' (e.g., distribution of copies, reader based, containing textual and visual information etc.) is also true of 'magazines'. Then whatever is true of 'magazines' is also true of 'consumer magazines', and so on. Since attributes of a given class are inherited by all its subclasses and sub-subclasses, all sub-subclasses are members of not only their immediate superclass but of every superclass above that one. Thus, if 'magazines' are a kind of publication, 'consumer magazines' are also a kind of publication. Chan (1999, 2003) reminds us that inheritance and transitivity are used for maintaining the overall coherent and logical structure (i.e., the generic relationship) of a hierarchical classification.

Mutual exclusion

An entity can belong to only one class in a hierarchy and the class can therefore belong to only one immediate superclass. This is called mutual exclusion. Although 'comic books' are like 'magazines' which have ongoing volumes published regularly, they cannot be classified in 'magazines' again if they have already classified in 'books'. This is also the reason for introducing cross-referencing when a class shares some attributes with other classes beyond its own hierarchy.

Prediction of association and distinction

A hierarchy is sometimes called an enumeration because the association and distinction rules of the hierarchy are predetermined. Without these rules, the hierarchy cannot be fully established. Thus, all entities in a class or all subclasses in a superclass are like each other in some predictable and predetermined way, and they differ from others in sibling classes in some predictable and predetermined way. In the example, 'magazines' and 'books' are alike in that they are both kinds of publication. They are differentiated from each other along some predefined criterion of distinction (in this case, 'magazines' are periodical publications but 'books' are not). Put in another way, if an entity or a class has the prescribed attributes of a class, it must be included in the class.

2.2.1.5 Rigidity is A Nature of Hierarchy

It has been found that the strict structural requirement of subject division in hierarchies makes them not always suitable for organising knowledge and for supporting user navigation (Ranganathan, 1951; Wynar & Taylor, 1992; Drengson, 1996; Kwasnik, 1999; Kim et al., 2006). In the book of “Introduction to cataloging and classification”, Wynars comments on a hierarchy:

"*that resulted in the assignment of fixed 'pigeonholes' for subjects that happened to be known or were foreseen when a system was designed but often left no room for future developments and made no provision for the expression of complex relationships and their subsequent retrieval.*" (Wynar, 1992)

In detail, the rigidity appeared in user navigation can be summarised in the following categories according to their relations to the characteristics of hierarchies.

Independent views of subject division

Class inclusion only defines that a hierarchy of classes (subjects) should be presented in a generic broad-to-narrow relationship based on the attributes carried by classes on each level. However, an entity is understood to have several, perhaps overlapping but separate sets of attributes and relationships, depending on the context and goal of the representation. It indicates that hierarchies could be different when representing the same subject based on their points of view. Table 2.1 clearly shows that Yahoo! Directory and the Open Directory view the Web from a more traditional perspective of knowledge organisation than Lycos while Look Smart and MSN tend to view the Web from the perspective of users' activities and interests.

Web Directory	Top Level Categories	Total No.
Yahoo! Directory	Arts & Humanities, Business & Economy, Computers & Internet, Education, Entertainment, Government, Health, News & Media, Recreation & Sports, Reference, Regional, Science, Social Science, Society & Culture	14
Open Directory	Arts, Business, Computers, Games, Health, Home, Kids and Teens, News, Recreation, Reference, Regional, Science, Shopping, Society, Sports, World	16
Lycos	Arts Entertainment and Games; Business and Finance; Careers; Computing, Internet and Mobile; Education and Society; Erotica; Gifts, Cards and Flowers; Lifestyle; Motoring; Music and MP3; Property; Homes and Gardens; Regional; Science; Shopping; Sports; Travel and Holidays	16
Look Smart	Auto; Cities; Education; Food; Health; Home Living; Money; Music; Recreation; Sports; Style; Tech & Games; Travel	13
MSN	Career, Family & Lifestyle; Education & Research; Money & Finance; Communicate; Entertainment; News & Sports; Online Safety & Security; Downloads, Services & Tools; Shop; Travel	9

Table 2.1 The Top-Level Categories of Popular Web Directories.

Retrieved August 19, 2008 from http://dir.yahoo.com for Yahoo! Directory; http://www.dmoz.org for Open Directory; http://directory.lycos.co.uk/ for Lycos; http://search.looksmart.com/ for LookSmart & http://specials.msn.com/ for MSN

This difference of view could be also presented even when dividing the same subject. Table 2.2 shows that the "science" subject is more traditionally perceived in the Open Directory than it is in the Yahoo! Directory.

The Subcategories of Science (Alphabetical)	
Yahoo! Directory	**Open Directory**
Aeronautics and Aerospace (191)	Agriculture (3,530)
Agriculture (2019)	Anomalies and Alternative Science (480)
Animals, Insects, and Pets@	Astronomy (3,993)
Anthropology and Archaeology@	Biology (27,119)
Artificial Life (59)	Chemistry (4,430)
Astronomy (3032)	Computer Science@ (2,193)
Biology (20375)	Earth Sciences (5,679)
Chemistry (1322)	Environment (6,223)
Cognitive Science (77)	Math (11,086)
Complex Systems (20)	Physics (4,652)
Computer Science (1370)	Science in Society (722)
Earth Sciences (4640)	Social Sciences (22,655)
Ecology (1091)	Technology (12,060)
Energy (590)	
Engineering (3686)	
Forensic Science (121)	
Geography (4388)	
Geology and Geophysics@	
Hydrology@	
Information Technology (87)	
Life Sciences (20)	
Mathematics (1714)	
Medicine@	
Meteorology@	
Nanotechnology (73)	
Oceanography@	
Paleontology@	
Physics (1836)	
Psychology@	
Space (1688)	

Table 2.2 The Comparison of "Science" Categories

Retrieved August 19, 2008 from http://dir.yahoo.com/Science for Yahoo! Directory & http://www.dmoz.org/Science/ for Open Directory

Nevertheless, in terms of user navigation, different views presented on the top-level classes of hierarchical classifications seem more important than the views on any other levels due to the top-down aspect of hierarchies which means they could be only accessed from the top. Olson (1998) states that a universal hierarchical classification is not always universally accepted because people may disagree on which 'sameness' is important and which should be used to categorise things on the top level of hierarchy. Thus, when a user does not agree with the view of division presented by a Web directory from observing its top-level classes, he would get in trouble with deciding a start point. Moreover, when a user does not agree with the division of a certain class in the Web directory, he would still have a problem in continuing his browsing. For example, the Open Directory contains 24 categories of wine but how can a user tell the difference between “Recreation: Food: Drink: Wine”, “Shopping: Food: Beverages: Wine” and “Business: Food and Related Products: Beverages: Wine” especially when they all have entries of wine makers, traders and sellers?

Division depth could be another problem of the “free views” of subject division as there are some practical limits before a hierarchy becomes too complex. Consider the placement of entries in “Regional: Europe: United Kingdom: Business and Economy: Shopping: Stores” of the Open Directory where most popular stores are listed. There are fashion chains like Clarks and Next, catalogue chains such as Littlewoods and Argos, supermarket chains like Tesco and Sainsbury and department stores like House of Fraser and John Lewis. Are they the same entities when considering their selling points? Note a hierarchy is not well designed to accommodate distinctions made along two very different sets of criteria. It is theoretically possible to further divide these stores by their types for a more objective or even more efficient classification but consider the current level of “Stores” in the directory and remember it is already sitting on the sixth level. Such further divisions would make the whole representation become too cumbersome and repetitive. Kwasnik (1999) reminds that if a hierarchy is weighted down by too many perspectives and disparate rules for grouping and differentiation, it loses some of its power as a clear representation. In user navigation, if a hierarchy is too simple, it is less informative and superficial but if it is too complex, it loses the advantage of being a clear and systematic view of the domain. Since multiple and diverse criteria are used by different hierarchical classifications, to do both simultaneously is representationally difficult.

Strict rules for class inclusion

Entities are included into a class of a hierarchy only if they possess all the necessary attributes defined by the class. At the same time, they are in only if they are sufficient to represent the class. This strict necessary-and-sufficient criterion is the nature of class inclusion. Thus, in a good hierarchy each member of a class is as good a representative of its class as any other. However, entities do not always conform to the necessary-and-sufficient criterion as people do not always perceive things so neatly. In reality, entities can belong to a class more or less with the result that one entity might be a better representative of a class than another entity because it fits the inclusion criterion better. For instance, the Open Directory places Amazon UK and John Lewis in “Regional: Europe: United Kingdom: Business and Economy: Shopping: Stores”, which implies that Amazon UK, Tesco, Boots and John Lewis are equally representative entities for the class “Stores”. To most British people, John Lewis is more representative than Amazon UK. Furthermore, entities in a class may share some attributes in common with each other, but not all might share the same attributes. John Lewis is a department store chain, Tesco is a supermarket chain and Boots is a pharmacy chain. Finally, an entity may belong to one class under some circumstances and to another class under other circumstances, or to both simultaneously. For example, Amazon.com is placed into “Shopping: Entertainment” of Open Directory but its UK company is placed into “Shopping: Stores <UK>” rather than “Shopping: Entertainment <UK>”. Without effective mechanisms “for indicating relative weight and presence of attributes and relative closeness or distance from some best-example prototype” (Kwasnik, 1999), the fuzziness and unbalance of class inclusion can easily cause misunderstanding. In user navigation, if this happens, a user may leave the category too early or too late when some top entries in the category are not that representative or even misleading.

Possibly incomplete and incomprehensive knowledge

Universal hierarchical classifications like the DDC and LCC attempt to be comprehensive because they are pre-established and enumerative. In order to achieve this requires relatively complete knowledge of the domain. This is not a problem with mature fields where the knowledge boundaries are clear. However, building a hierarchy for emerging fields seems not that simple as their theoretical frameworks have not been properly established. What would happen when such domains are rushed into a hierarchical classification? Kwasnik (1999) warns that such a representation which leads to premature closure in terms of knowledge creation would be misleading or skewed. Figure 2.2 shows a popular sign of premature closure – the "*Other* Media" in "Arts: Comics", which would apparently cause a problem in navigation as users may not be sure what it represents.

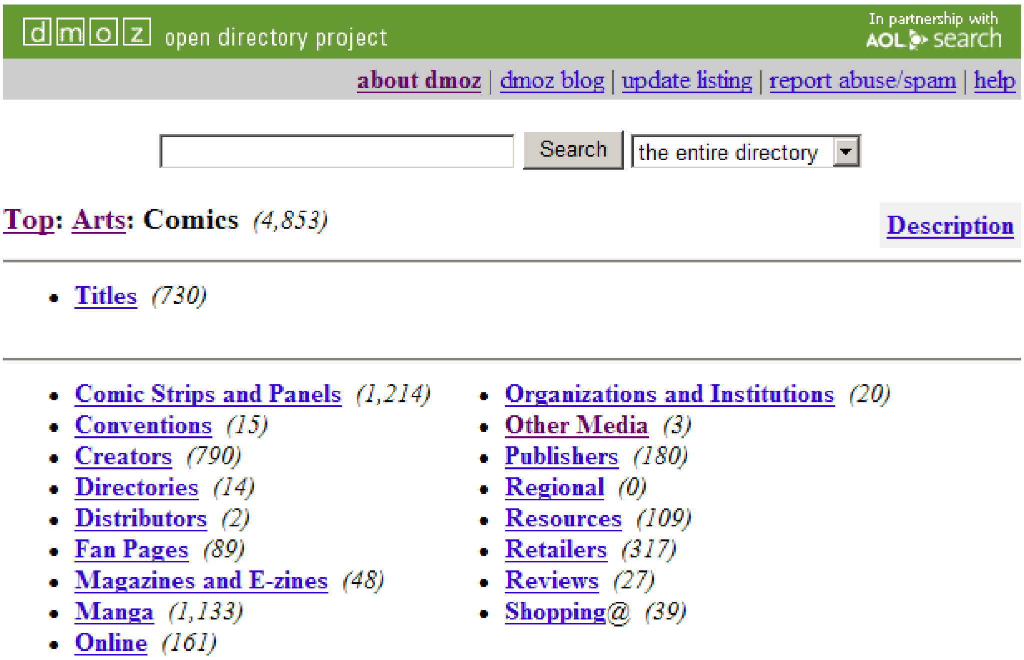


Figure 2.2 Category "Other Media" in the Open Directory

Retrieved August 25, 2007 from http://www.dmoz.org/Arts/Comics/

In addition, not only does the establishment of a hierarchical classification require complete and comprehensive knowledge of the domain in advance but also users need complete and comprehensive knowledge of the domain for using this classification. Kim (2006) asserts that users who do not have good knowledge of the domain have to make a series of trial-and-error attempts in locating subjects in a universal hierarchical system. This is because the success of using such a system lies in how well they understand the structure of the hierarchy and how exactly they know how the particular subject is classed under which one of the main classes (and divisions, subdivisions and so on).

Misleadingly cross-referencing

One important characteristic of hierarchies is that entities and classes in a hierarchy are mutual exclusive, which means that an entity can belong to only one class or a class can subordinate to only one immediate superclass. This property is not always practical because many subjects are understood to have several, perhaps overlapping but separate sets of attributes and relationships. For example, a comic book can belong to either magazines or books depending on the view of it but it can be placed into only one of them due to the compliance of hierarchies. The cross-referencing mechanism was introduced to bring certain degree of flexibility for accessing subjects without affecting the integrity of hierarchical structures. However, Koch & Day (1997) point out that only good cross-references can do the trick as cross-referencing can break down the original divisional chain in a hierarchy. Suppose a user is looking for a category containing website links to antique coin auction information in the Open Directory. First, he needs to "guess" broadly which top-level category could cover this topic because a hierarchical classification is a top-down (broad-narrow) system. Let us say he selected "Arts". Then, he needs to keep narrowing the subject level by level until he finds the right category. Note "Antiques" is a cross-referenced category in "Arts". So when the user follows it, he will be redirected to another divisional chain which is "Recreation: Antiques" from "Arts: Antiques". Then if he clicks "Auction", a cross-referenced category inside "Recreation: Antiques" again, he will jump to "Shopping: Auctions: Antiques and Collectibles". Then "Coins", then in "Shopping: Antiques and Collectibles: Coins: Auctions". Battles (2003) warns that the extensive use of cross-referencing will "make up an epistemological labyrinth unto themselves" because the divisional chain keeps changing instead of expanding.

Poor cross-cultural supports

A hierarchy does not support multiple views of a particular domain. Segall (1990) and Malt (1995) found that if the knowledge domain contains cross-cultural resources, it might not be suitable to build a hierarchical classification for satisfying culturally and linguistically diverse groups as the hierarchy will be somewhat ethnocentric (Olson, 1998). For instance, although the DDC is considered as the most frequently used scheme on the Internet (Koch, 1997, Williamson, 1997; Saeed & Chaudhry, 2001, 2002; Pollitt, 1998; Mitchell & Vizine-Goetz, 2002), Intner & Weihns (1996) make the criticism that the DDC reflects a "Western" outlook of the universe of knowledge which may not suitable for other languages than English. This point of view is supported by Kwasnik & Rubin (2003) whose experiment showed the difficulty in mapping and translating kinship terms in different classification schemes including the DDC and LCC to represent their concepts accurately in different cultures and languages. Moreover, Kim et al., (2006) compared the top-level categories and second level categories in twelve widely used commercial Web directories in four different Asian countries/regions and found significant differences in subject categorisation between American and Asian cultures. For example, in comparing local Yahoo directories and main Yahoo! Directory, "Humanities" is usually presented with "Arts" but was grouped together with "Social sciences" in Korea. "Politics" was added to the "Government" category in China and Hong Kong where it was placed under "Government" in the Yahoo! Directory. This means that such cultural differences of understanding would affect users in accessing subjects in a universal hierarchical classification such as the Open Directory or general Yahoo! Directory.

In summary, we can see that rigidity is the main cause of poor user navigation in Web directories and this comes with the strict structural requirement when establishing the hierarchy of Web directories.

2.2.2 Faceted Classification Schemes

Facet, French ***facette****, from Old French, diminutive of "****face****".*

2.2.2.1 Definition

Etymologically "facet" is one of the flat surfaces cut on a gemstone. The term was firstly introduced into Library and Information Science (LIS) by Ranganathan to describe his concept in the Colon Classification system (Maple, 1995). A universally accepted definition of facets in classification theory is given by Taylor (1992, 2000) as "clearly defined, mutually exclusive, and collectively exhaustive aspects, properties, or characteristics of a class or specific subject". Faceted classification schemes are also called *analytic-synthetic* schemes (Dykstra, 2004) as they break down each subject being classified in to its basic properties (analysis) and combine them to describe the subject content (synthesis). In other words, unlike hierarchical classification schemes which are top-down, faceted classification schemes are bottom-up systems.

Some people believe that the idea of facets was initially presented in the Dewey Decimal Classification (DDC) (Taylor, 2000) as it uses consistent mnemonics (notations) to notate categories regardless of their hierarchies. For example, 73 is used for referring to the US. on both sides of the decimal point such as 631.5973 for US. Cooking and 973 for US History. Apparently, 73 plays as a facet indicator for representing the specific regional information of any subjects (in this case, the US). Others give credit to the Universal Decimal Classification (UDC) for its widely used auxiliary signs which indicate various special aspects of subjects and relationships between subjects. For example, numbers following "=" in the UDC indicates the language (e.g., =20 means "in English" so 59=20 means Zoology described in English) and numbers following "+" means addition to (e.g., 59+636 means Zoology and animal breeding).

2.2.2.2 Classic Faceted Classification Schemes

The first fully implemented faceted classification system is generally acknowledged as Ranganathan's Colon Classification (CC) published in 1933 (Foskett, 1972; Kwansnik, 1999; Broughton, 2002). For Ranganathan, the problem with enumerative systems like the DDC and LCC schemes was that it could not enumerate a finite number of subjects to prescribe new areas of knowledge being discovered. In his book of "Philosophy of Library Classification", he wrote,

"*An enumerative scheme with a superficial foundation can be suitable and even economical for a closed system of knowledge…………What distinguishes the universe of current knowledge is that it is a dynamical continuum. It is ever growing; new branches may stem from any of its infinity of points at any time; they are unknowable at present. They can not therefore be enumerated here and now; nor can they be anticipated, their filiations can be determined only after they appear.*" (Ranganathan, 1951)

Ranganathan posited that any complex entity could be viewed from a number of perspectives or facets and postulated five fundamental elements which are Personality, Matter, Energy, Space and Time (PMEST) (Ranganathan, 1960, 1967):

(1) Personality: what the object is primarily "about". This is considered the "main facet", e.g., a ball.

(2) Matter: physical materials and abstract properties of the object. For example, the rubber of a ball, the ball's shape and colour.

(3) Energy: the processes or activities that take place in relation to the object. For example, bouncing a ball.

(4) Space: where the object happens or exists. For example, in a basket or storage room.

(5) Time: when the object occurs/happens.

He believes that any entity can be expressed in this so-called PMEST formula. For instance, if a document discussed "the design of glass vase in 19th century England", the facets would be as follows:

vase [Personality]; glass [Matter]; design [Energy]; England [Space]; 19th century [Time]

Thus, the strength of this classification comes through combining the pieces together to from the whole (Taylor, 1999). However, the Colon Classification did not gain wide adaptations like the DDC or LCC due to its regional constraints (Foskett, 2000) and the controversial PMEST fundamental facets (Miksa, 1998). Nevertheless, the facet theories behind the CC influenced a large number of classification schemes in the 20th century. For example, The British Catalogue of Music Classification for the British National Bibliography (Coates, 1960), the first edition of Bliss Classification (Bliss, 1953) and a later revised Bliss Bibliographic Classification (BC2). After noticing the significant theoretical success in classifying new objects (Kwasnik, 1992), Kwasnik (1999) summarised that "not all faceted classifications use Ranganathan's prescribed fundamental categories, but what they do have in common is the process of analysis". The process includes four steps:

(1) Choosing facets: forming the facets or fundamental categories;

(2) Developing/expanding facets: expanding and further developing these facets;

(3) Analysing entities: describing entities by using developed facets;

(4) Developing citation order: choosing the primary facet.

2.2.2.3 The Need for Facets

Ranganathan's motivation (1951) was to develop a classification scheme without a superficial foundation and is able to accommodate new knowledge. He claims that faceted systems could do the trick by combining prior existing categories (facets). Wynar also describes faceted classification in the book of '*Introduction to cataloging and classification (the 8th edition)*' as follows (1992):

"A faceted classification differs from a traditional one in that it does not assign fixed slots to subjects in sequence, but uses clearly defined, mutually exclusive, and collectively exhaustive aspects, properties, or characteristics of a class or specific subject. Such aspects, properties, or characteristics are called facets of a class or subject..."

Additionally in page 321, he says,

Faceted structure relieves a classification scheme from the procrustean bed of rigid hierarchical and excessively enumerative subdivision that resulted in the assignment of fixed "pigeonholes" for subjects that happened to be known or were foreseen when a system was designed but often left no room for future developments and made no provision for the expression of complex relationships and their subsequent retrieval.

And further in page 322,

"*... individual facets can be accessed and retrieved either alone or in any desired combination. This feature is especially important for computerized retrieval, which has been successfully applied to faceted classification, and in online retrieval as a complement to verbal retrieval by subject headings or keywords.*"

In detail, faceted classification schemes are specifically advantageous against hierarchical classification schemes in some aspects of knowledge organisation. These are illustrated by using the following example (Figure 2.3).

Type

Red wine | White wine | Rose wine | Champagne | Sparkling wine|...

Country/Region

Australia | Chile | France | Italy | New Zealand |...

Grape

Cabernet Sauvignon | Chardonnay | Merlot | Riesling | Sauvignon blanc |...

Price per bottle

Under £6 | £5 - £7 | £7 - £10 | £10 and over

Figure 2.3 The Wine Club Home of Tesco

Retrieved August 25, 2007 from http://www.tesco.com/winestore/

Hospitality

Facets are abstract attributes or properties of objects. Kwasnik (1999) claims that faceted classification schemes are able to accommodate new entities and facets as long as their fundamental facets are sound, which is particularly useful in classifying subjects where we have no way of predicting them or where the knowledge boundaries are not clear. If a new kind of wine produced in the future could be described by the fundamental categories of “Type”, “Region”, “Grape” and “Price per bottle”, the faceted classification scheme will be still robust.

Flexibility

Faceted classification schemes describe objects by assembling facets in an endlessly flexible way like building blocks to represent subjects (Vickery, 1960; He et al., 2003). For example, “all wine at less than £10 pounds” or “all wine made with Chardonnay at less than £10 pounds”. Kwanisk (1999) describes this feature as “post-coordination” in contrast to the “pre-coordination” required by hierarchies where attributes of a subject can be mixed and matched at the time of retrieval instead of setting them in advance.

Expressiveness

Faceted classification schemes can be very expressive as each facet is free to incorporate to other facets for representing any compound knowledge so does any multiple search (Lin, 2006). For example, with different combination of facets, it can make an expression of “all white wine from France at less than £10 pounds and made with a kind of grapes called chardonnay” or “all white and red wine from France and New Zealand at less than £6 pounds but not made with Chardonnay”.

Multiple perspectives

Facets are extracted from different and even overlapped views of objects so that faceted classification schemes allow users to view subjects with multiple perspectives. For instance, wine can be described as a kind of healthy substance, as a kind of drink, as a kind of product to buy and as an industry. Moreover, the support of multiple perspectives in faceted classification schemes also makes them possible to accommodate a variety of theoretical structures and models (Kwasnik, 1999; Denton, 2003). In facet-analysing a wine, one facet may reflect a particular model of wine types, another could be a model of source (grape) and so on.

Knowledge tolerance

Extracting facets from objects requires only a good understanding of the objects rather than comprehensive and complete knowledge of the whole subject domain. Kim et al., (2006) state that users can locate entities effectively and correctly as long as they understand the fundamental facets. Suppose a user is looking for a particular bottle of wine under 20 pounds. He does not need to have good knowledge about what aspects determine the price of wine (e.g., year, origin, grape, process and brand etc.). By combining the facets he knows, he can easily achieve the goal. For example, he starts with "Price per bottle" for the budget, then "Country/Region" for the place of origin and then probably "Type" for the basic types.

In summary, faceted classification schemes offer great flexibility and pragmatic appeal against the rigid hierarchical classification schemes in knowledge organisation.

2.2.3 Hierarchies versus Facets

In the recent years, the interests of using faceted classification schemes to replace hierarchical classification schemes in Web directories have been grown quickly (Duncan, 1989; Jones, 1990; Ellis & Vasconcelos, 1999, 2000; Ellis et al., 2000; Chan et al., 2001; Broughton, 2002; Bates, 2002; Patel, 2002; Denton, 2003; Adkisson, 2003;

Kim, 2006; Lin, 2006; Uddin & Janecek, 2007a, 2007b). However, whether hierarchical classification schemes or faceted classification schemes are suitable for Web directories should be determined by the role of Web directories rather than only considering their characteristics in knowledge organisation. To most users, the role of generic Web directories is to be a quality guide offering subject access to help users navigate and locate websites on the Web. This indicates that hierarchical classification schemes, at the present time, are still the first choice for Web directories in terms of the characteristics of good Web guides.

Knowledge coverage

A good Web guide does not have to include everything on the Web. Instead, it only needs to have a relatively good coverage. This is easy to achieve with hierarchical classification schemes because the establishment rules of hierarchies such as the aggregation and distinction of entities and the necessary and sufficient criterion of inclusion must be clearly defined in advance. Although the boundary of the Web is not yet clear, a good knowledge of the current Web is still enough to generate a good representation of it. However, since facets are abstract properties of objects which can be extracted without comprehensive domain knowledge, some subjects on the Web could be omitted with faceted classification schemes.

Systematic view

A good Web guide should provide users a high-level bird-eye view and holistic perspective to help them understand its representation and use it. Hierarchical classification schemes always have such clear and visible views through their unified hierarchical structures based on their single perspective of knowledge. However, faceted classification schemes cannot provide such unified visualisations with facets (Lin, 2006) because each facet is extracted from different perspective of an object and plays as an isolated kingdom which make it hard to have insight of the meaningful relationships among them (Kwasnik, 1999). For example, music is commonly classified by artist, genre, country, composer, instrument and so on while it is hard to find the relationship between a particular genre and its typical instruments (e.g., drum, keyboard, electric guitar and bass for Rock & Roll) or between a particular country and its popular genre (e.g., Britain and Britpop).

Real definitions

Web guides provide subject access on the Web and it is clear that real definitions of subjects are more natural than abstract definitions in terms of general acceptance. Hierarchical classification schemes extensively use real definitions to describe classes. Consider the category of “Video Games” in the Open Directory. It is easy to understand that it is a kind of game that use a video screen based on our understanding of game. On the contrary, faceted classification schemes use abstract and artificial facets which make them not always helpful. Consider the “Product Category” in Dell UK's laptop selector (Figure 2.4). Users need to find out how Dell defines “XPS”, “Inspiron”, “XPSGaming” and “RED” as these are not part of our understanding.

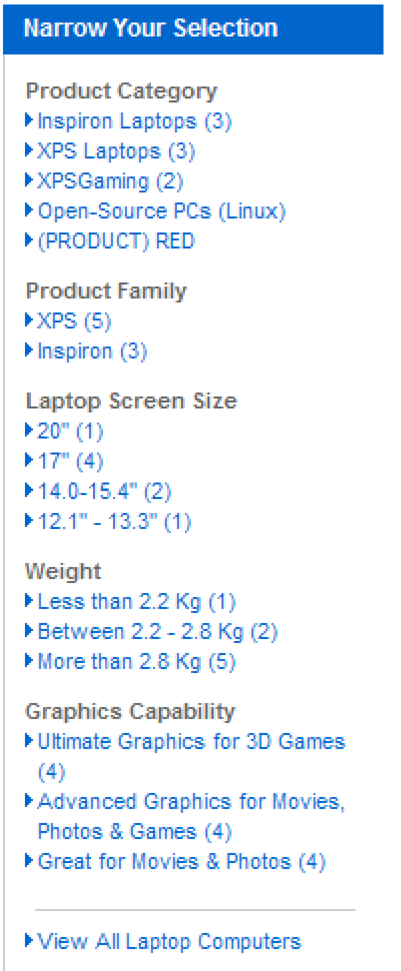


Figure 2.4 Dell UK's Laptop Selector for Home section

Retrieved August 25, 2007 from http://www1.euro.dell.com/content/products/category.aspx/notebooks?c=uk&cs=ukdhs1&l=en&s=dhs#subcats=xpsnb,laptop_studio,inspnnb&navla=&a=

Inference

Inference is very helpful for a Web directory as it allows reasoning from incomplete evidence so as to keep smooth navigation through categories. Hierarchical classification schemes support inference due to their strict class inclusion (i.e., transitivity and inheritance). Consider the category "Games: Video Games: Strategy: Real-Time: Tribal Rage" in the Open Directory. By observation and comparison with other video games, we could assess that Tribal Rage is a kind of real-time strategy game. We could even tell the difference between Tribal Rage and other strategic games in other categories although we do not know them. Faceted classification schemes do not support inference as facets are isolated from each other.

Matureness

Mature schemes can maximise not only the acceptance of Web directories but also their performance. Walt (1997) highlights the advantages of using hierarchical library classification schemes for organising Web resources including standardised thesauri, rigorous theoretical principles and rational structures. Drabenstott (1989) and Vizine-Goetz (1996) state the suitability of using the DDC and LCC to organise the Internet due to their widely accepted hierarchical schemes. Koch et al., (1997) reviewed major classification schemes used on the Internet and concluded that universal hierarchical classifications like the DDC was used more frequently than other schemes for the complete subject coverage, wide support, good familiarity and multilingual access. Williamson (1997), Saeed & Chaudhry (2001, 2002) and Chowdhury & Chowdhury (2004) found the DDC can provide better support for organisation of digital information resources with its latest edition. Jenkin et al., (1998) used the DDC to develop an automatic classification of Web resources for its comprehensive subject-based classes. Moreover, it was also found that the mature vocabularies and the similarity of subject headings (Iyer & Giguere, 1995; Olson & Ward, 1998; Koch & Vizine-Goetz, 1998;

Hiom, 1998; Chan, 2000; Qin & Stephen, 2001) used by hierarchical library classification schemes could be used to improve the hypertextual knowledge representation on the Web (Cochrane & Johnson, 1996; Pollitt, 1998; Koch, 2000; Hudon, 2000; Chan, 2000; Mitchell & Vizine-Goetz, 2001; Mason, 2008). Compared to hierarchical classification schemes, an obvious disadvantage of faceted classification schemes is the lack of universal schemes due to the difficulty in analysing universal knowledge (Kwasnik, 1999; Vickery, 1966; Tzitzikas et al., 2002; Lin, 2006). In addition, Koch (1997) argues that unlike home-grown hierarchical classification schemes which take advantage of mature hierarchical library schemes, the economic cost of maintaining self-devised faceted classification schemes will fall entirely on the originator of these home-grown schemes before they gain high popularities. Thus, Lin (2006) suggests that faceted analysis is more suitable for organising knowledge in relatively small and specific domains. This can be also used to explain that there are many approaches in developing universal faceted classification schemes (Ellis & Vasconcelos, 1999 & 2000; Broughton, 2002; Patel, 2002; Zins, 2002; Denton, 2003; Uddin, 2006; Uddin & Janecek, 2007a; Kim, 2006) but none of them has proved entirely universality and satisfactory.

Popular topics

A Web guide or directory should allow users to access popular topics as a basic requirement of subject access. Hierarchical classification schemes are top-down systems so they can easily present this feature by highlighting popular classes at each level. For example, the Open Directory (Figure 2.5) highlights some secondary categories such as “Movies”, “Television” and “Music” for the top-level category “Arts”. In comparison, since faceted classification schemes are bottom-down systems which mean that topics are actually “hiding” in the expressions made by various combination of facets, they cannot provide such quick access to popular topics.

Figure 2.5 The Homepage of the Open Directory

Retrieved August 25, 2007 from http://www.dmoz.org/

In summary, consider the essential characteristics of general Web directories in terms of good Web guides, hierarchical classification schemes are more appropriate than faceted classification schemes for Web directories.

2.3 Summary

Hierarchical classification schemes are most enormously used schemes in generic Web directories but their rigidity severely affects user navigation in their knowledge representation. This has given rise to an interest in using faceted classification schemes for Web directories. After comparing them for the role of Web directories, hierarchical classification schemes seem to be more competent than faceted classification schemes. The reasons are, first, hierarchical classification schemes offer relatively complete and comprehensive knowledge coverage in respect of their predetermination of establishment rules. Second, hierarchical classification schemes present a systematic view of the Web through their one-dimensional class definitions. Third, hierarchical classification schemes support real definitions and inference which could help user understand Web directories and keep their navigation smooth. Fourth, hierarchical classification schemes have library classification schemes as mature “backups” which maximise their acceptance and performance. Finally, hierarchical classification schemes are directive by allowing to access popular topics on the Web.

On the other hand, rigidity is the main cause of poor user navigation in Web directories and it comes with the strict structural requirement for knowledge organisation. That is, it will exist as soon as the hierarchies of Web directories are established. The next chapter conducts an usability inspection on representative Web directories for studying how rigidity is presented on the knowledge representation in terms of user navigation with the aim of discovering possible solutions.

Chapter 3 Misfits: A User's View of Rigidity

"Usability Inspection is the generic name for a set of methods that are all based on having evaluators inspect a user interface. Typically, usability inspection is aimed at finding usability problems in a design." (Nielsen, 1995)

3.1 Introduction

Our background study on the two most widely used classification schemes (i.e., hierarchical and faceted classifications) has restated the suitability of using hierarchical classification schemes in Web directories despite their rigidity in knowledge organisation. We have also discovered that rigidity is an issue that arises naturally from hierarchical structures and it is the main cause of user navigation difficulties. In order to understand how rigidity is presented in the knowledge representation of Web directories so as to affect user navigation, we conducted usability inspection studies on Web directories in this chapter. Moreover, we also set a research direction for further improvement based on the findings from usability inspection studies.

3.2 OSM and Web Directories

3.2.1 An Overview of Usability Inspection Methods

Usability inspection methods are one of the three general types of usability evaluation methods, along with usability testing methods and usability inquiry methods, which are typically used for finding usability problems in a design (Mack & Montaniz, 1994; Nielsen, 1995). These kinds of method mainly differs from user-based evaluation methods such as usability testing or inquiry in participatory design (Dumas & Redish, 1993; Wright & Monk, 1991). In user-based methods, usability problems are found through the observation of and interaction with users while they use or comment on an interface. In usability inspection, problems are found through the expertise of the inspectors and the inspection technique they use (Zhang et al., 1999). Compared to the other two types, usability inspection methods are particularly advantageous for their wide applicability. For example, in addition to apply usability inspections on completed interfaces, Nielsen & Philips (1993) point out that some inspections can be used for addressing issues like the severity of the usability problems and the overall usability of an entire design. Nielsen (1990 & 1992) also found that many usability inspection methods could be used to inspect user interface specifications that have not necessarily been implemented yet, which implies their suitability for the earlier stage in the usability engineering life cycle. Another key strength of usability inspection methods against usability testing and inquiry methods is they are informal to conduct (e.g., based on rules of thumb and the general skill and experience of the evaluators), easy to use and highly cost-effective (Jeffries et al., 1991). This is also why such methods are commonly known as "discount usability engineering" solutions (Nielsen, 1989 & 1993). It is common to divide them into two categories, practical inspection methods and theoretical inspection methods, based on the different level of interaction they concern. The former focuses on representational usability issues by understanding what users need to do for achieving a goal and the latter is interested in structural usability issues by understanding what users need to know before achieving a goal.

3.2.2 Practical Methods

Nielsen & Mack (1994) listed most practical usability inspection methods in their book of '*Usability Inspection Methods*' as follows:

Heuristic Evaluation: a version of usability inspection where usability specialists judge whether each element of a user interface follows established usability principles (Nielson, 1994; Nielsen & Molich, 1990).

Cognitive Walkthroughs: a review technique where expert evaluators construct task scenarios from a specification or early prototype and then role-play the part of a user working with that interface - "walking through" the interface (Polson, et al., 1992; Rowley & Rhoades, 1992; Spencer, 2000; Wharton, et al., 1994).

Formal Usability Inspections: a walkthrough method adapted from software inspection methodology where inspectors are formed from those involved in the product design for running walkthrough tasks to reveal encountering defects (Freedman & Gerald, 1990; Kahn & Prail, 1994; Gilb & Graham, 1993; Wheeler, 1996).

Pluralistic Walkthroughs: a group version of the walkthrough method where users, developers, and usability professionals step through a task scenario, discussing and evaluating each element of interaction based on their expertise as end users, developers or usability professionals (Bias, 1991, 1994).

Feature Inspection: a scenario based method where experts analyses the feature set of a product based on end user scenarios (Bell, 1992).

Consistency Inspection: a technique where an inspection team decides the different design elements of a product so as to ensure consistency across multiple products from the same development effort (Wixon, et al., 1994; Nielson, 1995).

Standards Inspection: a standardisation check where a usability professional with extensive knowledge of the industry standards analyses the elements of the product so as to ensure compliance with industry standards (Wixon, et al., 1994; Nielson, 1995).

Guideline checklists*:* listed by some researchers (Wixon, et. al., 1994; Nielson, 1995) for its conjunctive use with other inspection methods as a set of “expert” guidelines to judge the attributes and interaction methods of the product's interface (Hom, 1998).

3.2.3 Theoretical Methods

Theory-based usability inspection methods, which are also called user modelling based inspection methods, include:

Perspective-based Usability Inspection (PUI): a method that divides the large variety of usability issues along different perspectives based on an extended HCI model from Norman's “Seven Stages of Actions” (Norman, 1988) and focuses each inspection session on one perspective (Zhang et al., 1999).

Usability Pattern based Inspection (UPI): a similar method to Perspective-based Usability Inspection that defines domain specific patterns along with general usability collections and runs a single evaluation session for each defined pattern (Schmettow, 2005).

Ontological Sketch Modelling (OSM): a model adapted from ERMIA (Green & Benyon, 1996) and PUM (Blandford & Young, 1996) which identifies misfits between the designers' views of a product, the product itself and those of its users through the observations of entities, attributes, actions and relationships (Blandford & Green, 1997). The model has now evolved into CASSM (Concept-based Analysis of Surface and Structural Misfits) (Blandford et al., 2005).

3.2.4 Why OSM?

OSM was chosen from the rich pool of usability inspection methods for conducting our usability study on Web directories for three main reasons. First, OSM aims to discover structural usability problems rather than representational problems of a system. The definition of OSM is "*a structured but informal representation of the ontology – essential underlying structure – of a system, forming a basis for usability assessment*" (Blandford & Green, 1997). Hence unlike most usability inspection methods designed for spotting representational usability problems by understanding what the user needs to do (user tasks) on the system, OSM is interested in what the user needs to know rather than what they need to do. The intention of OSM is to yield a deeper understanding of the basic cause (e.g., the fundamental structure of a system) rather than the surface cause (e.g., some elements of the system) of usability issues. This matches our requirement for discovering the understanding gap between a hierarchical knowledge representation itself and the user's knowledge of the representation in terms of rigidity in knowledge organisation. Second, OSM is less dependent on the expertise and experience of inspector and is usable by non-specialists with good performance compared to practical inspection methods like heuristics and walkthroughs which are heavily reliant on the craft skill of inspectors (Nielson, 1994; Connell et al., 2002, 2003; Zhang et al., 1999; Blandford & Green, 1997). Third, OSM is suitable for working with an existing product interface as it aims to analyse the misfits among user, device and domain whereas most other methods are recommended for use in the early to middle stages of the product life cycle (Nielson, 1990&1992; Hom, 1998).

3.2.5 The Ontological Sketch Modelling (OSM)

3.2.5.1 Outline of OSM

OSM provides a common representation that supports reasoning about users, domains and devices. The representation is done by describing the entities, attributes, actions and any inter-relationships between them, that a user needs to work with when using a system. An ***entity*** is a 'thing' that a user has to know about and it may be relevant to the domain or the device or both. In the example of Microsoft Word illustrated by Blandford & Green (1997, 1998, 2001), 'character', 'word' and 'paragraph' are entities that are relevant to both the knowledge domain (word processing) and the device (Microsoft Word program). An entity also has some ***attributes*** that a user may change but not create or delete. For example, the 'font', 'size', 'color' and 'style' are attributes of a word. There are also some ***actions*** involved in creating or deleting entities, or changing attributes. For example, 'pressing <enter>' to start a new paragraph in Microsoft Word or 'highlighting' some sentences of a paragraph by moving the mouse pointer. ***Relationships*** are used to describe the relation between two or more entities, attributes and actions by identifying several types of connection between them such as 'consists-of', 'constrains', 'affects' and 'others'. Potential misfits are identified via the following entity-attribute typology:

User-private: entities that are part of the user's domain knowledge, but not directly represented in the device. For example, a user cannot compile free-ordered pages by using Microsoft Word like when they write in paper as Word can only produce fixed-ordered pages. These entities are likely to result misfits because they either cannot be represented by the device or have to be re-conceptualised by the user.

Device-private: device entities that a user has to know about, but cannot change easily and may not be able to see. One example is the Word style sheet although some style buttons (bold, italic, underline, align to left and bullets etc.) are available on a standard toolbar. These entities are hard for the user to learn.

Shared/device: entities that are explicitly represented in the device but not part of user's domain knowledge. For example, the file types (.txt - plain text, .rtf - rich text, .dot - document template etc.) of a Word file which are not a direct part of a user's document-creation domain knowledge. Although some of these entities may need to be learnt by novice users, once users get familiar with them, they become part of user knowledge. In this case, shared/device entities are unlikely to lead misfits.

Shared/domain: entities that are explicitly represented in the device, domain relevant and known to the user. For example, a 'word' or 'sentence' in Word. As with shared/device to arise misfits are unlikely.

In general, a potential misfit can be identified in either a *user-private* or *device-private* entity because they are not shareable or transferable to each other.

3.2.5.2 Method

We employed a similar methodology originally demonstrated by Connell et al. (2002) on the basis that of a similar knowledge domain (i.e., classification) and representation (i.e., digital library). We picked Open Directory in the period between March 2003 and June 2003, for its size (e.g., over 590,000 categories and 4,593,821 sites) and wide applicability (e.g., powers several other major directories including AOL, Lycos and Google).

An in-depth inspection of the Open Directory was carried out to identify the entities, attributes, actions and relationships embodied in the directory and hence to detect any potential misfits between device and users. The inspection had followed with two stages:

First, the main device entities and attributes were identified by inspection (as summarised in Figure 3.2). The main device entities consist of an extensive directory of website links, classified into 16 top level categories in which each has several tens of thousands of hierarchically linked sub-categories. Website links can be queried by using user defined keywords through a built-in search engine and then displayed as search results. In the Open Directory, a website link showing in a category points to the external Web addresses.

Next, the main user (Web directory users) entities (as summarised in Figure 3.3) were described from an initial analysis and a further user interview. The initial analysis identified a set of external links to be the content of a category and the search results via an internal search engine. An interesting feature of the internal search engine of the Open Directory is that the search results include both categories and websites relevant to the user's keywords. Detailed properties of these user entities were later discovered from the user interview.

Since categories are varied in their content due to the different levels in the hierarchy (e.g., top category, an ordinary parent, an ordinary child or an end category), it is also necessary to further describe the different types of categories as attributes of that entity. Then the relationship between, and properties of, the device and user entities were further identified.

In order to identify user entities, interviews were conducted with 5 potential users. The small number of subjects was used as the consideration of cost-efficiency (Virzi, 1992; Nielson, 1994 & 2000). The aim of the interviews was to identify any differences between the ways that the device (in this case, the Open Directory) manipulated entities (website links) and the typical use that a user makes of to the Web directory. It was, therefore, necessary to employee representative users to run the test. Generally speaking, users should have knowledge from three domains, the Web (as Web directories are Web applications), knowledge classification (as directories classify Web resources with certain principles) and general IT skills (for solving problems when interact with computer systems). We conducted an email interview to potential participants in order to discover their experience with typical services provided in these domains including questions like “please rate how familiar you feel to use library/the Web/computers for locating references/browsing online information/using software applications? (Scale 1 - 10)” in addition to their background check using questions like “please describe your professionalism”. After collecting responses from potential participants, we picked up five interviewees for their representative domain knowledge combinations as shown in Table 3.1.

- Interviewee 1 was an experienced librarian with good IT skills.
- Interviewee 2 was a college student with some IT skills who was also a regular library user.
- Interviewee 3 was an experienced IT help desk supporter with basic library knowledge.
- Interviewee 4 was a college student with very little IT skills and library experience.
- Interviewee 5 was a college student with no solid IT background or library experience.

Interviewee	**Related domain knowledge on a scale of *1-10***		
	General IT knowledge	**Library experience**	**Web usage**
1	●●●●●●●○○○	●●●●●●●●●●	●●●●●●○○○○
2	●●●●●○○○○○	●●●●●●●○○○	●●●●●●●○○○
3	●●●●●●●●●●	●●○○○○○○○○	●●●●●●●●○○
4	●●●○○○○○○○	●●●●○○○○○○	●●●●●○○○○○
5	●○○○○○○○○○	●○○○○○○○○○	●●●●●●●●●●

Table 3.1 The Domain Knowledge Profile of the Five Interviewees

The interview sessions were structured with ten questions designed to assess the extent to which the interviewees' knowledge and experience matched the requirements of the domain (here the general classification knowledge) and the device (here the Web directory). These questions were ranged from the general (e.g., 'What is a generic Web directory?') to the specific (e.g., 'How can you identify the contents of the ODP from a set of screen shots?') as listed in Figure 3.1.

The list of questions

Q1: Please describe what a generic Web directory is and how you make use of it.

Q2: Please describe what a category in a Web directory is about and what it presents for.

Q3: Can you look at these categories and tell their differences?

Q4: Please describe how websites are classified in a category of a Web directory.

Q5: If you were using a Web directory yourself, how would you expect to be able to access the topics in which you are interested (suppose that you did not have any particular interests)?

Q6: If you were using a Web directory yourself, how would you expect to be able to access the topics in which you are interested (suppose that you did have some particular interests)?

Q7: Please describe when you would use the internal search engine and what results you would except for.

Q8: Please state any difference between an internal search engine and a Web search engine (for example, Google).

Q9: Can you look at these supplied screen shots from the Open Directory and tell me if you recognise and understand the terminology?

Q10: Please try to outline a possible hierarchy to classify Amazon UK by looking at the homepage of the Open Directory and then find it out in the directory to justify your thought.

Figure 3.1 Questions Used in the Case Study of the Open Directory

We then used a think-aloud protocol to gather their answers with as much details as possible and decode them in the form of entities, attributes and relationships. For example, in Q2 “Please describe what a category in a Web directory is about and what it presents for”, an interviewee answered: “I think a category is a basic unit of Web directory which describes a subject of its parent and contains other categories presenting detailed sub-subjects of it. It may contains online resources too if there are any”. They were asked to check the accuracy of the record of their answers in the end of each question.

3.2.5.3 Results and Discussions

The main device and user entities of the OSM analysis of Open Directory are summarised in the form of tables as shown in Table 3.2 and Table 3.3 respectively where two entities: "category" and "keyword" were identified as private entities on both device and user sides.

Entity	Type	Description	Notes
Web Directory	Shared/domain	A Web guide that helps users navigate web sites by browsing topic based classifications.	It does not have to be complete in terms of sites but it has to be comprehensive in knowledge coverage.

Attributes	Instance	Notes
Type of directory	Open Directory is a directory that guides user to navigate on the Web	An internal directory could be a product directory eBay or Amazon used for helping their users browse classified selling goods.
Number of top level categories	Open Directory has 12 top level categories	The number of top level categories is decided and controlled by a web directory itself. So the number is a variable.
Number of categories	Open Directory has over 590,000 categories	The number of categories is decided and controlled by a web directory itself. So the number is a variable.
Type of classification	Hierarchical and topic based. Top: Business: E-Commerce	Categories are classified hierarchically and topic based. This is the basic rule followed by all web directories.

Entity	Type	Description	Notes
Category	Device-private	A category in a web directory. May contain sub-groups. May contain only website links. Searchable (for keywords relevancy search). Browsable by title (A-Z)	A category should follow the specific constructing taxonomy of a web directory.

Attributes	Instance	Notes
Name	E-Commerce	The name of a category is always short and concise, so users sometimes get confused by the meaning. For example, in *Shopping: Publications: Books*, there is a sub-category called "*General Interest*"

Attributes	Instance	Notes
		where all major book retailers are listed. All of our interviewees got confused when they were trying to identify where Amazon is.
No. Resources (if any)	E-Commerce (1,060)	The number presents how many web resources are collected under this category. However, most of our Interviewees said this is not really useful.
Classification hierarchy	Top: Business: E-Commerce	Although the name of a category may use some popular words representing online interests, users have to learn and follow the classification the web directory used in order to get useful information.
Direct sub-categories (if any)	Top: Business: E-Commerce: Strategy	The sub-categories belonging to a category. Available in all categories except end categories. May be difficult to predict.
Cross-referred sub-categories (if any)	Legal Information @	Sub-categories with '@' that are cross-linked to a category, but they do not belong to the category. Available in all categories except end categories. May be difficult to understand.

Entity	Type	Description	Notes
URL	Shared/Domain	A HTTP identifier pointed to a specific web resource	URL is the only way that users access information on the Web

Attributes	Instance	Notes
Title	eBay	The title of this linked resource when it is considered as a single document.
Summary	International person to person auction site, with products sorted into categories.	The summary of this linked resource when it is considered as a document.
Recommendation		Whether the resource is recommended by the directory
Order	By title.	Categories and websites are ordered by title.

Entity	Type	Description	Notes
Search engine	Shared/device	A tool that allows user to search	Preference: unclear (be default it will

Entity	Type	Description	Notes
		through the Web directory. By default, it will return either categories or web sites as results by default.	display both category results and website results but normally user would like to find out a specific category that matches their interest)

Attributes	Instance	Notes
Range	The entire directory, only within the current category.	By default, it will search the entire directory
Results filtering	Categories only, websites only and categories and websites (default).	The default setting is to display both relevant categories and websites.

Entity	Type	Description	Notes
Keyword	Device-private	The keywords allowed to perform a text search.	ODP allows very limited use of keywords due to its strict website naming and describing rules. For example, either “best travel agencies” or “toaster makers” will return no results on the ODP.

Attributes	Instance	Notes
Words or number of words	Any combinations users would like to use.	Again, ODP doesn't allow too natural queries.

Table 3.2 Open Directory's Main Device Entities and Attributes

Entity	Type	Description	Notes
Category	User-private	A category of a web directory. May contain sub-groups. Searchable (for keywords relevancy search). Browsable by title (A-Z)	A category should follow the specific taxonomy of a web directory

Attributes	Instance	Notes
Name	E-Commerce	The name of a category is always short and concise, so users sometimes get confused by the meaning. For example, in *Shopping: Publications: Books*, there is a sub-category called “*General Interest*” where all major book retailers are listed. All of our interviewees got confused when they were trying to identify where Amazon is.
Classification hierarchy	Top: Business: E-Commerce	Users have their own hierarchy preference when they use a

Attributes	Instance	Notes
		directory. So they may not follow the hierarchy the directory provided. They may jump from here to there, or follow the cross-referred categories to another hierarchy. All of these will affect the quality of their navigation.
Sub-categories (if any)	Top: Business: E-Commerce: Strategy	To most users, they think all categories within a category are the children of the category and they lack experiences in distinguishing direct children and cross-referred categories.

Entity	Type	Description	Notes
URL	Shared/Domain	A HTTP identifier pointed to a specific web resource	URL is the only way that users visit and explore the Web

Attributes	Instance	Notes
Title	eBay	The title of this linked resource after considering the content of it.
Summary	International person to person auction site, with products sorted into categories.	The summary of this linked resource after considering the content of it.

Entity	Type	Description	Notes
Keyword	User-private	The keywords that are used to perform a search task	User intends to use any words to describe their search demands but the ODP does not support a 'free-use' of keywords as much as a search engine does.

Attributes	Instance	Notes
Words or number of words	Any combinations users would like to use	variable

Table 3.3 Open Directory's Main User Entities and Attributes

Entity: category

The entity “Category” was identified as a device-private entity of Open Directory and also a user-private entity of its users mainly because the directory uses a “homegrown” classification scheme which does not always represent a user's view. In other words, since the scheme has its own classification “ontology” (Koch, 1997) for determining categories and their relationships, this ontology may not be fully perceived and/or accepted by users especially when they do not have a clear understanding about the different types of categories listed below.

Top level categories

Top level categories are root categories of a Web directory which presents a certain overall perspective of the Web and determines major hierarchies of the directory. These categories are also known as the starting point of browsing due to the fact that a hierarchical classification always moves from the general to the specific. Thus, top level categories are particularly important in user navigation as they indicate from which the main hierarchies start. For example, the 16 top level categories of Open Directory not only imply that the directory "thinks" that the whole Web can be viewed in 16 domains but also indicate that users need to agree with this arrangement and choose one domain to start browsing (Figure 3.2).

Figure 3.2 The Homepage of the Open Directory

Retrieved August 25, 2007 from http://www.dmoz.org/

None of our interviewees reported difficulties in understanding this 16-domain arrangement but one interviewee stated that he would prefer a Lycos-style arrangement as shown in Figure 3.3 as he thought it looked more practical and easier to follow. This indicates that a user's view of classification is sometimes determined by their activities on the Web rather than their real understanding on the knowledge domain as Lycos' Homepage classifies the online resources based on the popularity of searched subjects. In other words, to a certain extent, it would cause browsing problems if the domain is pre-established with a more advanced form of classification than that which the user can understand. Sometimes, if the user is specialised in certain knowledge domains (e.g., shopping), they may have no difficulties in the navigation no matter how these domains are viewed and classified in Web directories. However, as Web directories normally present a relatively general knowledge domain, these users may still encounter problems in several top-level categories before finding what they are interested in.

Figure 3.3 The Homepage of Lycos UK Directory

Retrieved August 25, 2007 from http://directory.lycos.co.uk/

Ordinary categories

Ordinary categories are any categories between the secondary level categories and the last level parent categories in a main hierarchy. These categories always have subcategories and sometimes contain a list of entities (URLs) as shown in Figure 3.4. At most time, they are transiting categories but they can also be a target category if they contain entities.

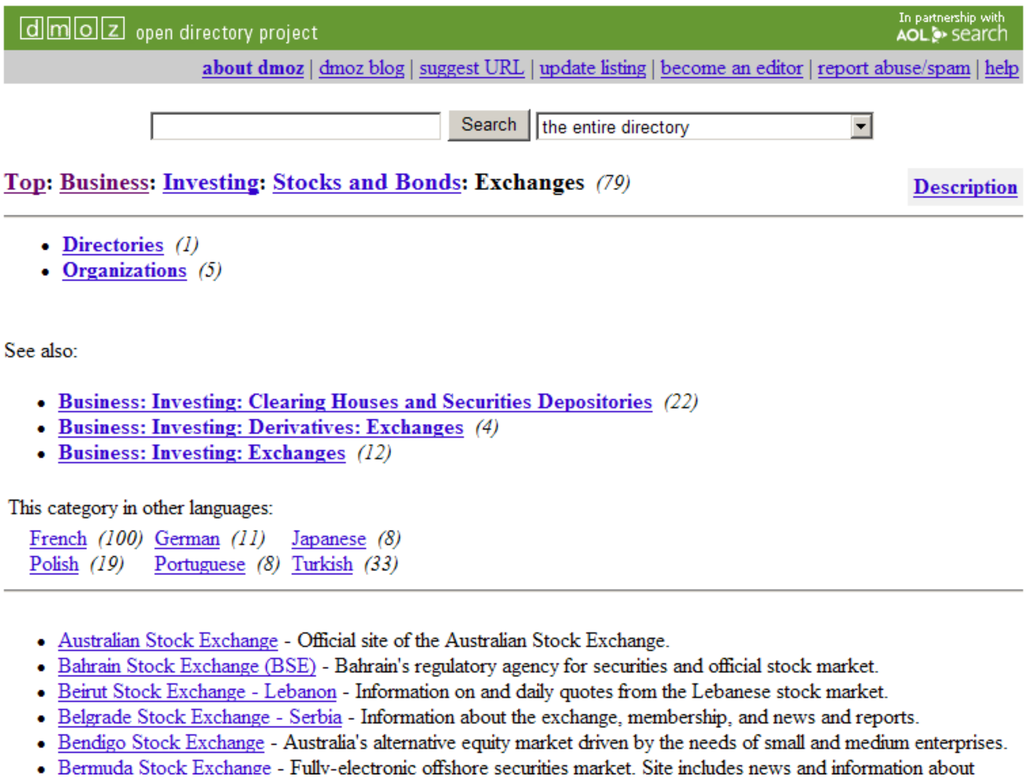


Figure 3.4 Ordinary Category "Exchanges" in the Open Directory

Retrieved August 25, 2007 from

http://www.dmoz.org/Business/Investing/Stocks_and_Bonds/Exchanges/

The real problem came out with these categories as we found most interviewees got stuck on certain levels of them and were unable to proceed with their browsing after choosing a top level category. From our observation, this situation was firstly triggered by the controlled vocabulary (thesaurus) used in the Open Directory, in other words, the definition and properties of a category set by the directory. Normally, users tend to map a term with the meaning they know without checking the real description of the term being used for a category. For instance, question 10 asked interviewees to find a category in the Open Directory which would contain Amazon UK and the correct answer should be "Regional: Europe: United Kingdom: Business and Economy: Shopping: Stores". Most interviewees failed to locate it correctly as they thought "Stores" referred only to physical stores (geographically distributed). They thought Amazon UK should be included in some categories like "online stores" or "online retailers". However, the term "store" used in Open Directory is for describing all distance and non-distance shopping instances (see below):

This category contains two types of sites.

A) Sites for retail chains who do not offer distance shopping but have bricks and mortar shops in more than one UK country, e.g., England and Scotland.

B) Stores offering a wide range of goods which do not fit into a subcategory within UK Shopping and allow shoppers to buy online, by mail, by phone, or by some other form of distance shopping.

Moreover, the wrong expectations in specifying an ordinary category with a certain general relationship also caused users to suspend their browsing as a hierarchical classification only allows one generic relationship between two categories. For example, books can be divided into either different subjects (e.g., fiction books, art books and language books etc.) in one hierarchy or different business natures (e.g., professional books, educational books and popular books etc.) in another hierarchy but only one in the hierarchy. Since to choose a further dividing rule depends not only on practical convenience but also individual preference, the type of hierarchy used by a given category and its subcategories is not always predictable by the user. For instance, most interviewees said they would have expected a similar division as the one used in the Yahoo! Directory to appear in Open Directory (Table 3.4) when they were asked for locating Amazon UK as a book seller. Instead, Open Directory divides book sellers in the basis of topics although we understand that both "Books" categories talk about book sellers from their general classification purpose (i.e., shopping).

	Yahoo! Directory	**Open Directory**
Classification	Business and Economy >	Shopping: Publications: Books

	Shopping and Services > Books	
No. of Sub-categories	11	49
Details of Sub-categories (first 5 categories)	Accessories (31) Book Search Services (52) Bookbinding and Conservation (37) Bookstores (6201) Business to Business@	General Interest (177) Antiques and Collecting (38) Arts (244) Audio (56) Biographies (39)

Table 3.4 The Top 5 Sub-categories of "Books" in Yahoo! Directory and Open Directory

Retrieved August 25, 2007 from

http://www.dmoz.org/Shopping/Publications/Books/

http://dir.yahoo.com/Business_and_Economy/Shopping_and_Services/Books/

Furthermore, cross-referred categories (subcategories following with the symbol "@" in any given category) are also problematic for continuous browsing in a hierarchy as a cross-referred category is not a direct child category of the parent category. That is, cross-referencing can harm the consistency of the hierarchical structure and the coherence of user navigation because it redirects users to a new hierarchy from the current accepted one, which may confuse them. We noticed that this happened a few times after some interviewees followed cross-referred categories. For instance, one followed "Business: Business Law@" from the top level category "Business" in the Open Directory to be redirected to "Society: Law: Legal Information: Business and Corporate Law". Sine the new hierarchy derived from "Society" is clearly different from the old one from "Business", it was not surprising when the interviewee told "I don't think I can browse further down because I have lost the focus of my thought".

End categories

End categories are the last level categories in any hierarchy. They contain no subcategories but a list of URL entries which indicate "the termination" of hierarchies (Figure 3.5). They are considered as user's goal in most time.


dmoz open directory project
In partnership with AOL search
about dmoz | dmoz blog | suggest URL | update listing | become an editor | report abuse/spam | help
Search | the entire directory

Top: Home: Consumer Information: Price Comparisons: Books (32) Description

See also:

- **Shopping: Publications: Books: Used and Rare: Search, On-line** (16)

This category in other languages:
German (8) Swedish (5)

- AAABookSearch.com - Compares prices at 44+ online bookstores.
- AddAll - Compares prices among 41 on-line bookstores, including used and out of print titles.
- AKABook - Offers price comparison and search for new and used books, textbooks and DVDs.
- AllBookstores.com - Compare prices at over 30 online bookstores by title, author, ISBN, or keyword.
- AllDiscountBooks.net - Compares prices from online retailers in US, UK, and Canada.
- Any Book For Less - Compares prices at 22 online booksellers, including delivery to US, Canada and UK.
- Best Book Deal - Compare prices at 47 online bookstores. Calculates tax and shipping cost.
- Best-of-the-World.com - Compares prices and shipping for new and used titles.
- BookButler - Compares prices at up to 40 stores.
- BookCost.com - Compares prices at several online retailers, and displays current coupons available.
- BookFinder.com - Book shopping search engine, plus a meta search of 12 directories of used, rare, and out of print booksellers.

Figure 3.5 End Category "Books" in the Open Directory

Retrieved August 25, 2007 from

http://www.dmoz.org/Home/Consumer_Information/Price_Comparisons/Books/

Concerns for the extent of subdivision were brought up with some end categories in Open Directory when two interviewees thought the site lists in some end categories need to be further divided into at least two sub-categories to offer better navigation support. For example, in the category of "Major Retailers" (Figure 3.6), all retailers either online or off-line in all trade categories including fashion, home, books, electronics and even superstores are listed here. When the list is long, it is not efficient to locate some of them with similar nature (e.g., all supermarkets) in an alphabetical order.

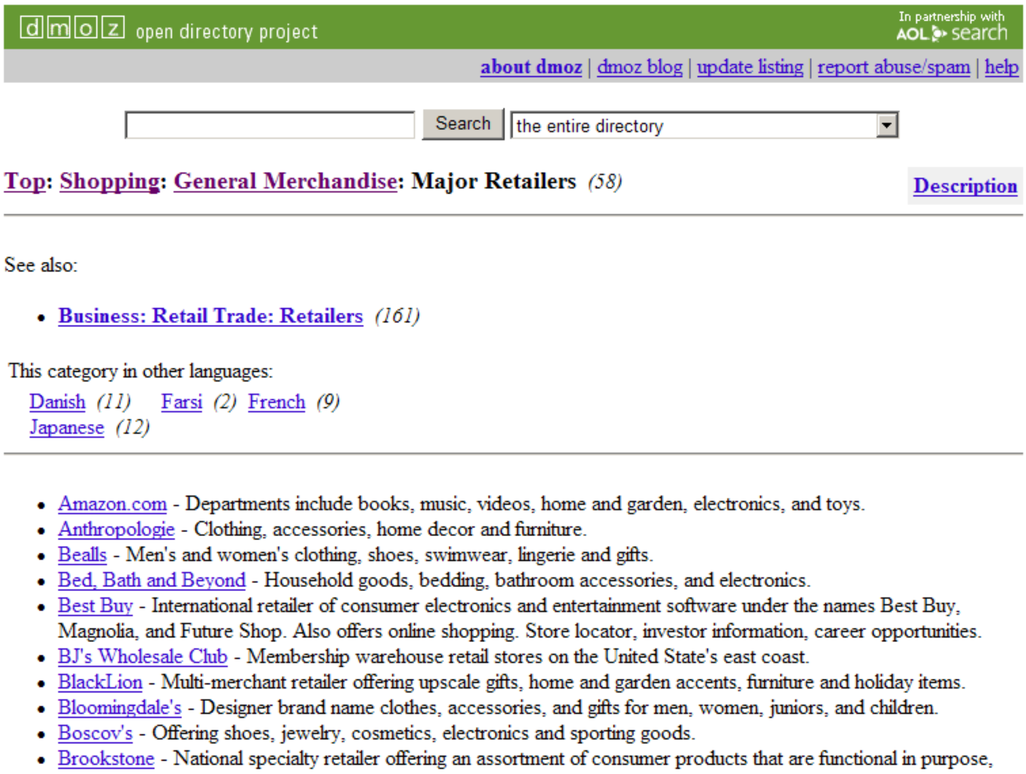


Figure 3.6 End Category "Major Retailers" in the Open Directory

Retrieved August 25, 2007 from http://www.dmoz.org/Shopping/General_Merchandise/Major_Retailers/

Moreover, we found that naming end categories with numbers and letters also caused navigation problems as some interviewees complained this kind of categorisation was not effective unless users have good knowledge of the subject. For example, in Figure 3.7, numbers and letters appeared as sub-ordinate categories under "Arts: Animation: Anime: Titles" indicating the category is sub-divided in an alphabetical order. That is, category "3"is for Anime shows whose title starts with the number 3 (e.g., 3x3 Eyes). This kind of classification is only useful when a user wants to find some information about a specific Anime. However, the most common use of a directory is to guide and inspire users, especially novice users to discover subjects they are interested in but not familiar with so it should not expect its users to have a good knowledge beforehand. In addition, the number following a category indicates how many resources are contained in this category and its sub-categories. It may be a good hint to indicate whether a user is on the right track of their journey but most interviewees said they would still continue even if the number is small when they thought they were in the right place.

Figure 3.7 Category "Titles" in the Open Directory

Retrieved August 25, 2007 from http://www.dmoz.org/

Entity: Keyword

Keyword is another entity which was identified as both device-private and user-private, which was mainly caused by the users' misunderstandings in using a Web directory's search engine. All interviewees sensed that Open Directory's internal search engine was just a tailored Web search engine which should work in the same way as Google, Yahoo! Search or Ask.com. The only difference, they said, is that it only searches for local directory information instead of the Web. This mistaken understanding of a search engine is common because of the huge impact of Web search engines have had. However, the fact is that a directory's search engine differs from a global search engine in many aspects due to the different purpose of the search. Web directories classify websites into hierarchically organised categories based on the sameness and distinctiveness predefined by these directories. This grouping concept indicates that a

Web directory is used for finding specific categories containing a group of websites sharing the same properties rather than particular websites. Thus, users are not expected to use the internal search engine in the same way as a Web search engine. That is, when they are thinking of relevant keywords to describe their interests, they need to think of the relevant keywords which might be used to describe a potential category. One problem here is, the depth or extent of classification can be unclear to users so they may use too specific keywords. For example, if a user is searching for “toaster retailer”, the Open Directory would return “no results found” (Figure 3.8) as such a category is too detailed. In fact, the directory only has broader categories like “appliance retailers” or “electronics retailers”.

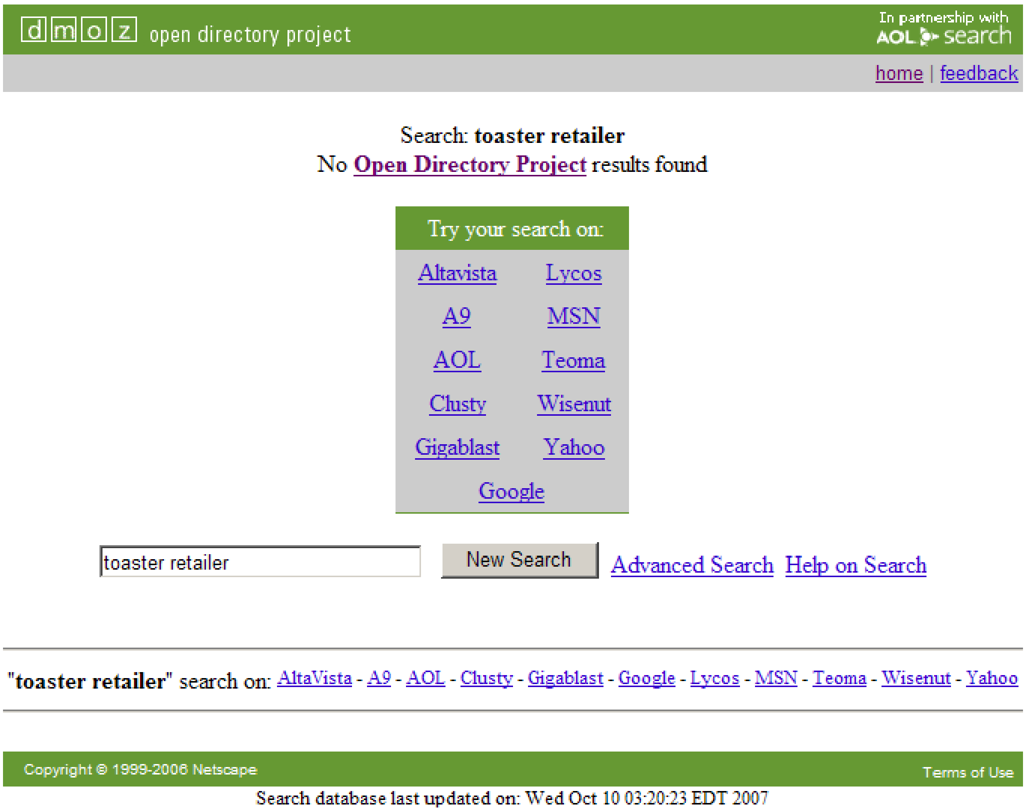


Figure 3.8 The Search Results for “toaster retailer ” in the Open Directory

Retrieved August 25, 2007 from http://www.dmoz.org/

Another problem is that keywords allowed to form a query to get satisfactory results from an internal search engine is more constrained than it is from a Web search engine. This is due to the strict rules used for category and site descriptions in order to maintain the consistency and quality of content. This aspect is aggravated further when a directory tries to maintain an unbiased attitude of its classification. For example, Open

Directory suggests "*descriptions of sites should describe the content of the site concisely and accurately. They should not be promotional in nature*" (Open Directory, 2008). More strictly, Yahoo! Directory regulates that "*description suggestions must be no longer than 25 words and refrain from using any marketing language or slogans*". For the title, it says "*make sure not to suggest a title longer than five (5) words and if your site is commercial, the title submitted must be the company name*" (Yahoo, 2008). Such rules and principles make the content of a category accurate, dispassionate and concise but also make it more artificial compared to the content of a normal page on the Web. Consequentially, even when a user understands the difference between a Web search engine and a directory's search engine, he may also fail in a search as the range of keywords that can be used are very limited. For example, queries like "cheapest mobile retailers", "UK book stores" or even "European automakers" submitted to Open Directory would return either no or unsatisfactory results.

3.2.6 Suggestions for Further Research

This OSM case study for evaluating Open Directory has identified two major misfits between the Open Directory and the user models of such a directory in relation to the domain of classification. These misfits are also common in other general Web directories with hierarchical classifications.

The first aspect of misfits is that the core concept of establishing classification schemes is still that of peer review although most home-grown schemes used in Web directories can be related to some well-known library classification scheme such as the DDC (Dewey Decimal Classification) or LCC (Library Congress of Classification). This indicates that a user will still have difficulties in understanding the ways in which the Web is perceived, categories are derived and organised in a Web directory. When they cannot obtain a clear understanding of a Web directory, their own views would become dominate during their navigations. In this case, the rigidity of hierarchies could be easily amplified so as to affect the user's further navigation. Thus, from a user's perspective, the key of minimising misfits between they and Web directories lies in whether they could establish a correct understanding of these directories. That is, the misfits would be improved as long as their understanding is improved. Information Visusualisation (InfoVis) and Web Persoanlisation are both popular domains dealing with the improvement of user understanding of information representation.

Researchers in Information Visualisation generally consider the misfits of understanding as a consequence of non-distortion-oriented techniques used for representing large hierarchies (Monk et al., 1988; Beard & Walker, 1990; Donelson, 1978; Herot et al., 1980; Leung, 1989). The non-distortion-oriented approach provides all the information at the same detail level so it can only display a portion of the information at a time due to the constrains of display devices (Clementi, 2007). If there is a large hierarchy, it means that users have to scroll and use paging to access to the remainder of the hierarchy. Websites are commonly organised and represented in this way. For example, Open Directory composes of many hierarchically linked web pages and each webpage actually represents a category of the directory. A user can only access one webpage at a time on their screen. If this webpage is longer than the actual display area, the user has to scroll down the page in order to access more information on the page. Alternatively, they can also click at a link of the page to access another category. The major weakness of this kind of technique is that the information displayed in a static page leaks of context, which makes its interpretation difficult. In this case, the user cannot obtain a good global view of the directory until he completes visiting all the categories of the directory. To address this problem, distortion-oriented techniques (focus + context), which utilise transformation and magnification functions to allow the co-existence of local details with global context at the same time, were introduced (Leung & Apperley, 1994). That is, a user's focus of information will be displayed with great detail on a section of the screen, while the remaining information is rendered with less detail at the same time but it is still kept on the screen to provide an overall context to facilitate navigation (Stasko et al., 2000; Stasko & Zhang, 2000). Typical distortion-oriented approaches for visualising large hierarchies include hyperbolic tree (Phillips & Gunn, 1992; Gunn, 1992; Munzner & Burchard, 1995; Lamping et al., 1996), Treemaps

(Johnson & Shneiderman, 1991; Johnson, 1992; Jungmeister & Turo, 1992; Turo, 1994; Bederson et al., 2002), Botanical Visualisation (Kleiberg et al., 2001), Cheops (Beaudoin et al., 1996), cone tree (Robertson et al., 1991; Carriere & Kazman 1995), MoireTrees (Mohammadi-Aragh & Jankun-Kelly, 2005), Fractal Trees (Koike & Yoshihara, 1993; Ong et al., 2005), TreeJuxtaposer (Munzner et al., 2003), FlexTree (Song et al., 2004) and Reconfigurable Disc Trees (RDT) (Jeong & Pang, 1998) etc.

On the other hand, researchers in Web personalisation believe that the misfits of understanding between users and a website are mainly generated from the unsorted content the website always contains. That is, even if a Web directory has a good representation, it could still cause a user understanding difficulties as long as it is not exclusively designed for the user. This is because the user always needs to put great effort in filtering relevant content. Thus, approaches from this area put emphasis on tailoring the content of an information representation according to users' personal interests so as to make the representation easy to use.

Approaches in both these directions have their advantages and disadvantages but the thesis set its research direction for Web directories in Web personalisation as in the next chapter.

The second aspect of misfits is that it is unclear to users that the search engines for Web directories are mainly used for locating categories rather than searching particular websites like what a Web search engine does. This suggests the necessity of redefining the search model for Web directories which will be covered in the next chapter after the discussion of Web personalisation techniques.

3.3 Summary

We have applied Ontological Sketch Modelling to evaluate the Open Directory and identified two typical misfits mainly caused by the users' inadequate understandings of Web directories in terms of the conceptual model of their hierarchical classification schemes. Taking the Open Directory as an example, these findings restate that rigidity is the main cause of user navigation difficulties in Web directories. Therefore, for the first misfit, we set our further research direction to Web personalisation and for the second misfit, we planned to redesign the search engine model for Web directories which are both discussed in the next chapter.

Chapter 4 A Unified Framework for Improving Navigation

"*Web Personalization can be defined as any action that makes the Web experience of a user personalized to the user's taste. The experience can be something as casual as browsing the Web or as (economically) significant as trading stocks or purchasing a car." (Mobasher et al., 2000a)*

4.1 What is Web Personalisation?

Personalisation was originally a marketing term commonly referred as *one-to-one* marketing (Riecken, 2000), which involves a process of tailoring a product or service to an individual user based on their personal characteristics or preferences. The aim is to improve a user's experience of a product or service in a way that is exclusively designed for the user. A common definition of Web personalisation is "any action that tailors the Web experience to a particular user, or set of users" (Mobasher et al., 2000a). To some extent, any Web browsing activity that aims for enhancing an individual user's experiences in browsing, navigation and search on the Web can be seen as a Web personalisation approach. However, this should not be confused with *customisation* which occurs when a user is able to configure an interface with some preferred options (e.g., changing the number of results displayed per page from 10 to 100 in Google's search results setting). Nielson (1998) and Bonett (2001) argue that if the control of the look and/or content is explicit and is user-active, it is customisation; on the other hand, if the control is implicit and user-passive or at least somewhat less user-driven, it is personalisation. This way of distinguishing customisation and personalisation by considering user involvement is somewhat vague as users can still be actively involved in certain processes of personalisation (e.g., user profiling). Thus, distinguishing whether it is personalisation or customisation roots on two aspects. First, if the user specified information is delivered through explicit system functions, it is customisation. For example, a user is able to display and hide some content presented on the interface. If it is delivered via implicit system analysis, it is personalisation. For instance, the system predicts a user's interest by using certain rules for observing the user's activities and delivers content on the basis of analysis. Second, if the user specified information only concerns some changes of the look of a system or is an explicit part of the original content, it is customisation. For example, a user chooses to display specific sections on the homepage of a website. On the other hand, if it is re-processed content based on the original content in terms of the user's characteristics, it is personalisation. For example, a user comments and rates some of his favourite songs on a website, and then the website recommends some songs from the repository based on his ratings.

Web personalisation is now a hot topic due to e-commerce's serious “push” where companies seek to build better customer relationships and more profitable websites through tailored services. So the e-commerce industry also has a more practical understanding for the concept. For example, “*personalisation refers to a feature that allows providers of online products and services to make use of information about their customers to interact with them on an individual basis, for instance, in providing specific types of information or in cross-selling products*” (Bossard, 2001). The following example - Amazon's personalised user's store (Figure 4.1) is a typical personalisation service which use a user's records including the past purchase (e.g., “items you own”), search and page view (e.g., “Page You Made”) to generate individual recommendations.

Figure 4.1 [Username]'s STORE at Amazon.co.uk

Retrieved August 25, 2007 from http://www.amazon.co.uk/ (with cookies enabled)

In summary, Web personalisation offers a user-centred navigation experience on a general-purpose information representation on the Web by delivering content based on the user's interest.

4.2 A Process-oriented View of Personalisation Techniques

Tailoring is a frequent verb used in the definition of personalisation, which actually describes a process that measures a user (for understanding his interests), cuts off unnecessary content (in which he is not interested) and then delivers the rest to the user. This strong process-oriented view has been widely accepted by researchers in presenting an architectural view of Web personalisation. For example, Mobasher et al., (2000a) outline a general architecture of automatic usage-based Web personalisation by dividing the overall process into two components: an *offline component* (batch process for data preparation and usage mining) and an *online component* (online process for recommendation). Pretschner & Gauch (1999) describe personalisation system as two main processes: *user profile creation and representation* and *content filtering/rating* (e.g., collaborative/individual filtering). Thomson (2005) proposes a standard cross-site framework for Web personalisation involving two main processes: *client side identification and personal data storage* for summarising personalised data and *server side personalisation* for generalising personalised page. Adomavicius & Tuzhilin (2005) claim that personalisation constitutes an iterative cycle of *Understand-Deliver-Measure* process in which each consists of two stages: data collection & profiling (Understand); matchmaking & delivery (*Deliver*) and impact measuring & strategy adjusting (*Measure*).

Web personalisation is generally composed of two main processes: *user profiling* and *content filtering* (Figure 4.2). User profiling, which normally happens at the client side, is a process that involves implicit and/or explicit data collecting (user actions, browsing histories and other usage data streams etc.) and interpreting from users following with using learning algorithm for modelling user interests. Content filtering, which is always a server side process taking place after user profiling, is a process that rates and filters in-site and/or cross-site content based on what user profiles present processed by filtering rules and then deliver the results back to users.

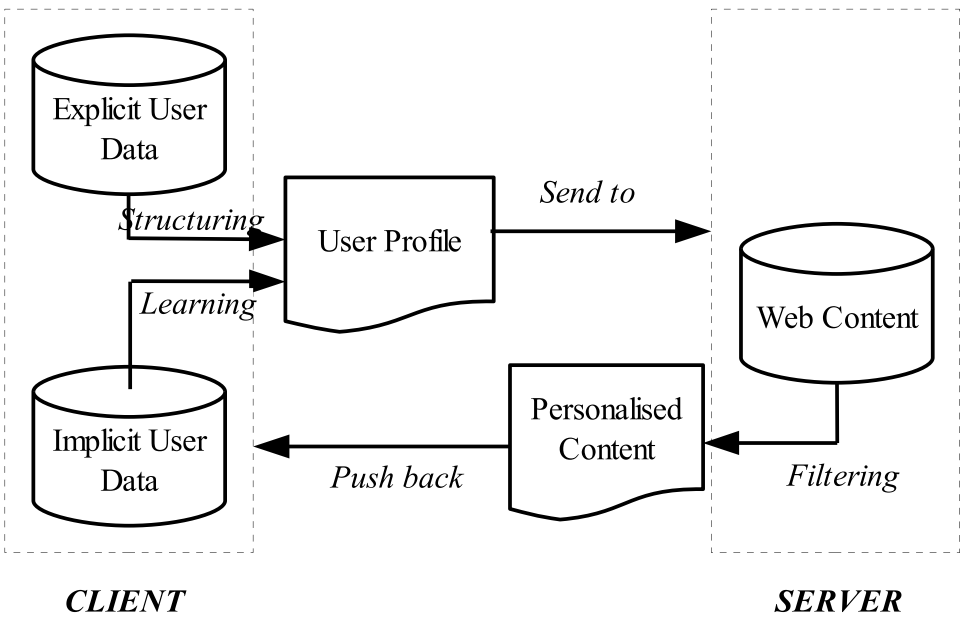


Figure 4.2 A Process-Oriented View of Web Personalisation Architecture

4.2.1 User Profiling Techniques

A user profile is a systematic view of a user showing the user's interests, preferences or information needs. It is generally considered as the key of Web personalisation which gives the system “the *ability* to deliver experiences” to the user (Bonett, 2001). The process of taking such a picture of a user either *explicitly* or *implicitly* (Rich, 1998; Thomson; 2005) is called *user profiling* or *user modelling*. Explicitly users tell the system about their interests in a mutual understandable way. For example, a user subscribes for some topics he might be interested on a website by checking or un-checking a range of topics provided by the website. Implicitly the system learns their interests by studying their behaviour anonymously (the user would not be bothered). For instance, an e-commerce website makes further purchase suggestions to a user according to his past purchase records on the site. In this case, the user did not tell the system what he is interested in buying next but the system makes some guesses and recommends them to the user.

4.2.1.1 Explicit User Profile

An explicit user profile is built through active involvement of the user, typically through fill-in or tick-check forms. The most beneficial point of an explicit user profile is that users are actively involved in the information gathering process. They tell the system what they like so that this kind of profiling method provides the most accurate profile about the users. Moreover, once the profile is set up, the information is upfront so it does not require the same information to be rebuilt through repeated use. For example, if a user sets up his regular tube route updates through the Transport for London website, then the updates will be automatically sent to him via emails so that he does not need to set up the service every time he wants to check such information.

However, since users are actively involved in building their profiles, a possible drawback is that if they are reluctant to spend time on setting up, the service remains underused as the profiles are not completed (Nielson 1998; Manber et al., 2000). The issue might be enlarged when complicated forms or options are introduced for establishing detailed user profiles. It could also explain why many scholars suggest using explicit user profiles for addressing relatively simple navigation issues, i.e., personal information space (name-space) issues. For example, PAINT (Personalised, Adaptable Internet Navigation Tool) allows the user to organise a tree-like personal view of the Web by providing an interface for logging and categorising their visited locations (Oostendorp et al., 1994). The idea was then implemented commercially as Bookmarks used in Web browsers (e.g., Favourites in Internet Explorer and Bookmarks in Netscape). Personal portals such as My Yahoo!, My MSN and iGoogle can be seen as recent approaches in this direction. In addition, personalised content subscriptions (e.g., personalised forum topicsubscription) and content-targeted advertising (e.g., Hotmail customised advertising) are also well-established fields in this area.

Another downside to explicit user profiling, as argued by Mobasher et al (2000a) and Bonett (2001), is that an explicit profile may remains static after initial setup, which would cause its performance degrade over time as the profile ages (unless the user remembers to make constant updates in his profile). Considering the same example of personalising travel information updates, if the user moved to a new place without updating his regular routes, he will keep receiving travel updates corresponding to his outdated profile which does not represent his current needs.

4.2.1.2 Implicit User Profile

An implicit user profile does not require user's active input. Compared to explicit profiles, such profiles are established automatically and transparently through the system's observation, study and analysis of a user's past and/or present Web activities. In other words, users are only passively involved during the creation of user profiles – in most occasions the profiling process cannot be noticed or observed. The advantage of an implicit user profile is significant. First, it requires no extra work on the user side (Bonett, 2001) and second, it intuitively adapts to the change of user's interests accordingly (via continuous monitoring and learning). Moreover, since implicit user profiles are generated from a relatively comprehensive and intensive collection of user behaviours and Web usages, they allow more complicated and various personalised applications (e.g., recommender systems) based on them. For example, Amazon studies past user activities and makes product recommendations for a user such as “Customers who viewed this item also viewed...” or “Customers who bought this item also bought...”. Alexa (2008) also uses collective usage patterns (i.e., a user's surfing history) for assessing sites and determining related links.

However, an increasingly noticed disadvantage is that choosing data collecting and learning algorithms for building implicit user profiles usually depends on personal experiences and tastes due to the lack of user profiling standards/guidelines. Thus, the accuracy of profiles relies heavily on the algorithms and techniques employed during the process (e.g., link analysis, usage mining or URL clustering). For example, Mobasher et al., (2000a) emphasises that the experience of employing appropriate methods (e.g., choosing association rules, deciding parameters and thresholds used on them etc.) affects the precision of identifying user sessions and transactions so as to acquire the user profiles from raw usage data.

Sometimes the issues of user privacy and data protection in implicit data collecting could also be problematic (Volokh, 2000; Peppers & Rogers, 2001; Chellappa & Sin, 2005) as users may not want some activities (e.g., online banking, legal or medical consulting) and sensitive information (e.g., user account) to be exposed to other parties. Bonett (2001) suggests introducing user supervising mechanisms to show them what information is gathered and how that information is used or shared or to use privacy statements and data use policies when collecting user data. However, psychologically it may still make users uncomfortable if they know they are being monitored (Thomson, 2005). This issue seems to be more serious in cross-site than in-site profiling especially when user profiles are stored in the server or the user profiling agent acts as a proxy linking users to the Web.

4.2.1.3 Hybrid User Profile

There are also a few approaches combining the use of explicit and implicit data to construct user profiles. Normally, such profiling methods are considered as implicit profiling techniques for the involvement of the autonomous data collection process (Pretschner & Gauch, 1999). However, we suggest classifying them as explicit profiling processes if the user's inputs are compulsory for finalising the data representation of user profile. A common situation is, a user's data are gathered by an autonomous agent for initialising the user's profiles and then the user will be asked to check the accuracy of them to finalise the profile. For example, Syskill & Webert (Pazzani et al., 1996) construct a user profile by asking the user to review the current viewing Web page as *hot*, *lukewarm* and *cold* so as to learn whether he is interested in the page (link). Other profiling processes such as the ones used in FAB (Balabanovic, 1997), IfWeb (Asnicar & Tasso, 1997) and SiteIF (Stefani & Strapparava, 1998) require peer review for adapting user profiles created implicitly. We suggest call these approaches as *hybrid*. This is because although initial profiles are created by system, the constant updates for the profiles need to be done with user's feedbacks.

4.2.2 Profile Learning Techniques

User profiling is not only a process that decides what kind of user data is collected (e.g., users' input in a form or their Web usage) and how to collect them (e.g., explicitly asked or implicitly gathered) but also a process that learns and represent the user's interests from the data collection. The second part is not always necessary for explicit user profiling as the user's interests are clearly represented to the system through some shared channels (e.g., a form). For example, if a Hotmail user checks and ticks some marketing information that he wants to receive, he is actually “telling” the content provider what he likes in a way the provider also understands (i.e., a internally formatted tick-check form with invisible keywords and tags). However, such formatted user interests cannot be easily obtained in implicit user profiling as the user data collection contains a huge amount of rough and unsorted information in various formats. Hence, it requires some extra data interpretation work (e.g., log cleaning, user session identification, transaction identification, link analysis etc.) for correctly learning the user's interests. For instance, the user's surfing history in Web browsers is commonly used as the information repository for building user profiles. Such histories contain all web pages (links) the user has visited in a certain period. How would a user profiling system know which pages represent the user's interests if it cannot ask the user straight away? In one way, the system could measure various information about each page such as the number of hit counts of the page, the time spent on viewing the page or the keywords presented on the page etc. Then the system may think that the pages with more hits, longer viewing time or higher keyword similarities are what a user is interested. There comes another question: some of the user's regular activities can also contribute high hit counts and long viewing time on certain web pages which may normally not considered as a part of the user's interests. For example, web pages with high hits may come from a default homepage setting in the user's Web browser, an email account the user always checks or a forum he often visits. More complicated learning analysis will be required if user profiles are constructed from their usages (e.g., following a link, starting a new browsing session, shopping, emailing, etc.). For example, it is always difficult to determine whether two links are related to each other if the user has just jumped from one to the other without any explicit signs (e.g., following a link in the previous page). Therefore, choosing the right learning algorithms is a key for implicitly obtaining accurate and comprehensive user profiles.

Traditional user profiles that consist of simple factual information, for example, demographic data, are called *factual profiles* (Adomavicius & Tuzhilin, 2005). Factual profiles capture relatively simple and straight user data (e.g., website history logs and keywords) for producing certain facts about the user (e.g., user's favourite search is about *travel* and favourite websites are shopping sites). These profiles are usually represented by a set of weighted words or keyword vectors. According to Pretschner & Gauch (1999), learning algorithms for factual profiles mainly come from related Information Retrieval fields, typically from text learning and document (page) measurements.

4.2.2.1 Keyword Extraction (Factual Profiles)

A very popular approach from Information Retrieval (IR) is to describe the process of using a set of weighted (or representative) words to index and identify the content of a document. The technique has been extensively used to construct weighted keywords vector profiles from analysis of the content of web pages that user have visited, which aims for content filtering oriented personalisation. For example, if a user visited eBay.co.uk and Amazon.co.uk, his profile may look like an array of combined keywords extracted from relevant meta tags (e.g., *description* and *keywords*) of the two websites. In this case, "eBay = auction, fixed price, books, cars, computers, digital cameras, DIY, DVD, jewellery and music & Amazon = digital camera, LCD TV, books, DVD, low prices, video games, pc games, software, electronics, home, garden, video, amazon". Such keywords of a web page can be extracted from using a vector-space model (tf-idf) (Salton & McGill, 1983; Armstrong et al., 1997) and other alternative IR methods (Harman, 1995) after word stemming/weighting process (e.g., Porter Algorithm (Porter, 1980)). For example, FAB, an adaptive web page recommendation service, uses vector-space model and word stemming to generate a representation of 100 highest-weighted words for per user-visited web page as the basis of the user's profile (Balabanovic, 1995, 1997). Similar approaches also include WebWatcher (Armstrong et al., 1995; Joachims et al., 1997), Personal WebWatcher (Mladenic, 1996), Letizia/Let's Browse (Liberman, 1995, 1997, 1999), ifWeb (Asnicar & Tasso, 1997), WebMate (Chen & Sycara, 1998) and Web Personae (McGowan et al., 2002) etc.

Keyword extraction is sometimes used in conjunction with *relevance* feedback (Baeza-Yates & Ribeiro-Neto, 1999; Foltz & Dumais, 1992; Harman, 1995; Buckley & Sulton, 1995) for improving the accuracy of user profiles. For instance, FAB (Balabanovic, 1995, 1997) allows users to use a 7-point scale user rating system (explicit relevance feedback) to update their profiles. IfWeb (Asnicar & Tasso, 1997) and SiteIF (Stefani & Strapparava, 1998) extract data autonomously (implicit relevant feedback) from the documents on which the user explicitly expressed some (positive or negative) feedback to update and refine his profile. In addition, WBI shows a further derived approach from IR by clustering keywords extracted from user visited web pages into different categories of interests for offering an organised view of user interests (Barrett et al., 1997).

4.2.2.2 Page Measurements (Factual Profiles)

Page measurements are composed of a set of user action measurements for determining whether a web page is related to the user's interests or is a just a “pass-by” page. Common user activities accounting for measuring page relevance include page viewing, bookmarking, link following, page scrolling and mouse activities (i.e., left/right-clicking and pointer highlighting/moving). Such measurements can be done through statistical based quantitative and/or qualitative analysis (Goecks & Shavlik, 1999) and rule-based heuristics (Liberman, 1995, 1999; Mladenic, 1996). For example, Letizia determines whether a user is interested in a page by considering the idle time (reading time) on a page and its link following status (immediately return or spent some considerable time) (Liberman, 1995, 1999). The agent also excludes links on a page if they get naturally passed over by normal reading behavioural (from top to bottom and left to right). Goecks & Shavlik (1999) utilise a combination of measurement including hyperlinks clicked, scrolling activity and mouse activity for predicting user's interests in a page. SiteSeer (Rucker & Polanco, 1997) constructs a user profile from bookmarked links. In other words, it hypothesises that links collected in the bookmarks by a user are what the user is interested in. Page measurements are often used with keyword extractions for presenting a higher degree of understanding of user interests. For example, WebWatcher (Armstrong et al., 1995; Joachims et al., 1997) and Personal WebWatcher (Mladenic, 1996) learn the quality of a page from its keywords representation and linking status (i.e., whether it is a link found on a previous page user visited or not).

Profiles that capture more sophisticated user data from users by studying their Web usage are known as *behavioural profiles* (Adomavicius & Tuzhilin, 2005). Compared to factual profiles, behavioural profiles are usually used for more targeted and specified personalisations. For example, an e-commerce website discovers that many of its users follow the link from price comparison and product review websites to its website, then it will show up relevant review and comparison results in product pages to attract them. Adomavicius & Tuzhilin (2005) summarise commonly used modelling (learning) techniques for constructing behavioural profiles in three areas as follows.

4.2.2.3 Conjunctive Rules (Behavioural Profiles)

Conjunctive rules are association or classification rules used in profiles for providing an intuitive, declarative and modular way to describe user behaviour (Adomavicious & Tuzhilin, 1999, 2002). For example, a user always uses Orange241 promotional code obtained from the mobile operator before watching a film. This habit can be a part of the user's profile for describing his rule for watching films (Adomavicious & Tuzhilin, 2001). Such rules can be learned from the transactional history of the user using various data mining techniques (Hand et al., 2001; Schechter et al., 1998; Buchner & Mulvenna, 1998). For example, suppose a user watched *Jumper* at Cineworld West India Quay on Wednesday evening with Orange241 offer. Corresponding conjunctive rules built for the user could be something like MovieType = “Action and/or Adventure and/or Sci-Fi” > “When” = “Wednesday evening” > “Where” = “Cineworld West India Quay” & “Discount” = “Yes”. Then on the next Wednesday, Orange would automatically send a sms like: “We thought you may like *Hannah Montana 3D: Best of Both Worlds Concert Tour* and don't forget to get a 241 before visiting Cineworld West India Quay”. Here

Hannah Montana 3D: Best of Both Worlds Concert Tour is selected for its relevancy based on all the rules.

4.2.2.4 Web Browsing Sequences (Behavioural Profiles)

Web browsing sequences are a series of activities that a user typically performs in certain websites. For example, when a user visits the gadget website Firebox.com, they usually start their browsing from the homepage, then goes to "*tech toys*" section, then browse the "*work:play*" section and the "*experiences*" section next, and then leave the website. In other words, their regular activities on the website can be summarised as a sequence like "Firebox: homepage > tech boys > work:play > experiences > exit". Such sequences can be identified and learned from transactional histories of users using frequent episodes (Mannila et al., 1995, 1997) and various association rules (Agrawal & Srikant, 1995; Srikant & Agrawal, 1995; Cooley et al., 1999; Han et al., 1998; Han & Fu, 1995).

4.2.2.5 Signatures (Behavioural Profiles)

Signatures, which are also called evolving profiles, are significant entities that can be aggregated from large data streams of simple transactions over time (Cortes et al., 2000; Mobasher et al., 2000b). Acting like a trigger, signatures are commonly used for monitoring statistical significances in a user browsing history. For example, a signature inserted into a user's profile by Amazon.co.uk for later producing a tailored store experience could be "top 10 most frequently viewed product categories over the last 30 days". Then the user's transactional data of product categories in Amazon.co.uk which may consist of linking requests and linking time stamps would be used for analysing the top 10 most viewed products based on calculating total visiting times and length.

In summary, in addition to using collection methods to classify user profiles into *explicit* and *implicit*, user profiling approaches can be also classified into *simple* (factual) and *advanced* (behavioural) in terms of learning objectives. Moreover, the quality of factual user profiles rests squarely on keyword extraction mechanisms whereas the accuracy of behavioural user profiles relies heavily on the sequence and transaction identification of user data.

4.2.3 Content Filtering Techniques

Besides profiling, content filtering is another important process for personalisation, which uses matchmaking technologies to deliver targeted content and services for the users based on the information (i.e., user interests) represented in their profiles. Unlike profiling which consists of two sub processes (i.e., *data collection* and *learning & representation*) the core task of filtering is matchmaking. In other words, the process is "another crucial aspect of personalisation that depends on the quality of the underlying matchmaking technologies" (Admavicius & Tuzhilin, 2005). There are several ways to classify matchmaking technologies into broad categories. For example, Breese, Heckerman and Kadie (1998) present a technical and algorithmic-based view by dividing filtering techniques into *Heuristic-based* and *Model-based*. Pretschner and Gauch (1999) summarise filtering processes into *individual* and *collaborative* according to the user focus (i.e., one user or a community of users). Payne (2000) classifies known techniques for personalisation into *rules-based*, *collaborative filtering or community based* and *inference* from the applicability fact. Balabanovic & Shoham (1997) describe the approaches as *content-based*, *collaborative* and *hybrid* based on the recommendation approach – here we use their categorisation to classify matchmaking technologies as described below.

4.2.3.1 Content-based Filtering

Content-based filtering is a traditional technique for personalisation which has its root in the Information Retrieval (IR) community. The concept of content-based filtering is about tailoring Web pages (for their content), services and products based on what a user liked in the past. That is, a content-based filtering system selects items based on the correlation between their content and user profiles (preferences) (Van Meteren & Van Someren, 2000). Typically, Web items are treated as text documents and their content is represented with a number of weighted words extracted through weighting scheme. Then only items that have good similarity (keyword occurrences) with the user's profile (in terms of their weighted keyword representations) would get recommended. For instance, three top weighted keywords in a user profile for his eBay usage are “Agatha Christie” (*author*), “Pentax FA lens” (*Pentax autofocus lens type*) and “Alchemy Gothic” (*jeweller*). When a new product starts selling on eBay which is related to any of these keywords, it would be forwarded to the user. Since content-based methods analyse the content of Web items based their textual representations, statistical measurements (e.g., cosine similarity, n-grams), probabilistic user models and classification learning algorithms (e.g., Bayesian Classification, neutral networks, Nearest Neighbours, Decision Trees) have been extensively employed from IR domain.

Rule-based filtering which applies simple logical rules (i.e., if *this* then *that*) for delivering specialised content to user, can be seen as a primitive or limited kind of content-based filtering. For instance, when a user reviews the printer he added into his shopping cart for checkout, he will find promotion deals like “buy 100 sheets of photo paper and save 50%” or “add 2 colour inks for a free delivery” remain as unchecked items in the basket. The key to this scheme is that the developer must know ahead of time what the personalisation should be (e.g., promote overstocked items) so that he can develop relevant triggering rules (e.g., if adding a product to a basket and/or proceed to checkout then recommend bundled deals). Payne (2000) argues that this is strict and must constantly be evaluated and adjusted depending on the business' needs and the scale of this scheme can be also very large if more detailed personalisation is required. For example, when a user adds a Fuji F480 digital camera into his basket at

Jessops.com, the photographic retailer will return the user a list of related services, accessories and products in which some items have already checked by default as recommendations (Figure 4.3). This is good for users but for the online store, such rules of recommendations have to be done and such personalised items have to be decided for every single camera selling at the online store in advance.

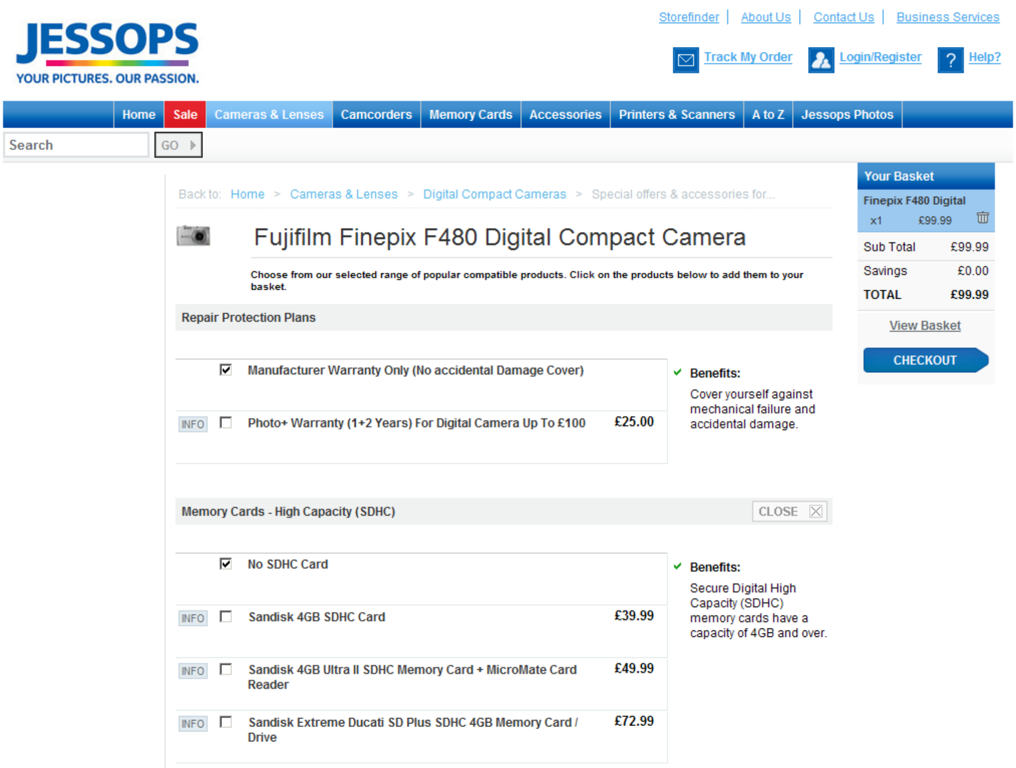


Figure 4.3 An Example of Rule-based Filtering in Jessops

Retrieved November 27, 2007 from http://www.jessops.com/Products/Configure.aspx?soid=67286&kitid=67286-1

A pure content-based filtering system is simple, fast and easy to implement but it has several shortcomings. First, textual representations of Web items come from a shallow analysis of certain kinds of content. Even for Web pages, the representations capture only certain aspects of the content and there are many others that would influence a user's experience. For example, IR techniques ignore aesthetic qualities of a Web page such as multimedia information (i.e., embedded images, video and audio), advertisements and network factors (e.g., loading time etc.). A second problem is personalisations or recommendations based on textual comparisons are over-specialisation. This is because only items scoring highly for the similarity against a user's profile will be recommended and this restricts the user from seeing other items similar to those already rated. This is often addressed by injecting a note of randomness (e.g., using crossover and mutation operations as part of a genetic algorithm (Sheth & Maes, 1993)) to increase the chance of seeing other items with low similarity scores. For example, many commercial websites such as Amazon, eBay or YouTube use a randomly sliding bar of product recommendations.

4.2.3.2 Collaborative Filtering

Collaborative filtering (CF) or community based filtering is an increasingly popular technique extensively used in the e-commerce industry nowadays. Rather than recommend items which are similar to items a user has liked in the past, items are recommended on the basis they are items other users who are similar to the user have liked. That is, a collaborative filtering system chooses items based on the correlation between people with similar preferences (user profiles). Basically, for each user a set of "nearest neighbour" users is determined with whose past activities show strong similarity to them. Then only items are found with good scores of interests in these nearest neighbours would be recommended to the user. For example, SiteSeer (Rucker & Polanco, 1997) suggests new URLs to a user based on discoveries of the user's virtual neighbours in terms of the similarity of bookmarks. That is, if a URL is found in two user's bookmarks, it will measure the degree of overlap of their bookmarks. Then if the overlap is high, which means there are many common URLs in their bookmarks and the two users can be considered as “virtual neighbour”, SiteSeer will recommend the rest of URLs from one's bookmark to another. Examples of systems taking this approach by determining “nearest neighbour” also include GroupLens (Resnick et al., 1994), FAB (Balabanovic, 1997), WebWatcher (Joachims et al., 1997) and WebACE (Han et al., 1998) etc.

Item-based collaborative filtering is a very common type of collaborative filtering seen in the commercial websites and popularised by Amazon.com. The approach only considers an item as a unique identifier without considering the similarities of users and then makes suggestions solely based on the relationships of other items linked to the item in various ways by users. For instance, Amazon tells the user “Customers Who Bought This Item Also Bought” when the user is browsing Jan Williams' book *Welcome to Britain: A Celebration of Real Life* (Figure 4.4).

Figure 4.4 An Example of Collaborative Filtering for the Book “Welcome to Britain” in Amazon.co.uk

Retrieved November 27, 2007 from http://www.amazon.co.uk/Welcome-Britain-Celebration-Real-Life/dp/0755314476/ref=pd_sim_b?ie=UTF8&qid=1196175316&sr=1-1

Collaborative recommendation is “free” from all the shortcomings mentioned for content-based systems but this approach also introduces certain problems of its own. One well known issue is the “First-Rater” problem (Wikipedia, 2008; Payne, 2000; Balabanovic & Shoham, 1997). That is, if a new item appears in the system, there is no way it can be recommended to a user until more ratings are obtained through other users' actions (e.g., reviewing, buying or specifying which other old items it is similar to). So if items are not popular or well-known, it becomes hard to recommend them. For instance, Amazon.com's “Customer Who Bought This Item Also Bought” or “Customer Who Viewed This Item Also Viewed” recommendations are always not available for unpopular items which are listed on the last few pages sorted by “best selling” option. Common solutions for this problem include running cross-promotions and bundled deals to artificially establish connection between popular items and unpopular ones and using featured recommendations to increase the popularity of these cold items. For example, Amazon's “Customer Who Bought Like This Also Bought” option. However, such solutions are not effective if the website contains a huge amount of dynamically changing information. For example, if items are removed from the database quickly, their relationship with other items will be emptied accordingly. In this case, the recommendation link between these items and other items will be invalid. This is also why systems like Amazon often offer a mixed-mode method of search including browsing and searching to reduce the impact of ineffective collaborative filtering results. Another problem is called “Cold-start” problem, which is often associated with collaborative filtering on the basis of the similarity of users. This is caused by new users in the system who have not submitted any ratings to let the system determine their preferences so as to make recommendations. The size and composition of the user population are also keys to user-based collaborative filtering. For example, if a user has unusual tastes compared to other users in the system, the user will receive poor recommendations for the lack of similar users in his group. Moreover, “if the number of users is small relative to the volume of information in the system, there is a danger of the coverage of ratings becoming very sparse, thinning the collection of recommendable items” (Balabanovic & Shoham, 1997).

In addition, a collaborative filtering system recommends items based on the *user ratings*, which is a general name of all relevant user activities indicating their interests, for example, rating, reviewing, buying or even frequently viewing an item. Since ratings are normally done without considering their content, the lack of access to the content of the items prevents similar users from being matched unless they have rated the exact same items. For example, if one user usually visits the BBC weather page for acquiring London's five-day weather forecast and another always checks MSN weather page for the same purpose, the two users would never become nearest neighbours although these pages contain the same content.

4.2.3.3 Hybrid Approaches

Hybrid approaches combine the use of collaborative and content-based methods, attempt to inherit “generic advantages” (Balabanovic & Shoham, 1997) and avoid disappointments from both sides. That is, they aim to solve two common scaling problems for all Web services (i.e., an increasing number of users and an increasing number of Web items) and enhance group awareness and communications at the same time. The combination can be achieved in two ways. One way is to implement content-based and collaborative filters separately but combine their results to produce the final recommendations. A popular instance is where Amazon shows “Customer who viewed (bought) this item also viewed (bought)...” on a product page a user is viewing (an implementation of collaborative filtering). At the same time, it also creates a tailored user store link on the top navigation bar showing recommended products (an application of content-based filtering). The other approach, as implemented in Fab – an adaptive, multi-agent system for recommending Web pages (Balabanovic, 1997), is to develop both a content-based agent and a collaborative agent in a single recommendation model to generate situation-driven recommendations. For example, if the prerequisite of triggering the collaborative agent cannot be achieved (i.e., if an item remains unseen by others), Fab will recommend items through the content-based agent. On the contrary, it will make collaborative recommendations if the current content analysis of the item is incomplete and imprecise. Moreover, if both conditions can be fulfilled, Fab will recommend items to users in a combination style.

In summary, content-based and collaborative filtering techniques aim for different personalisation purposes whereas hybrid approaches are a combination trying to provide recommendations for both purposes. Adomavicius & Tuzhilin (2005) classify them into *simple* and *advanced* based on their performance comparison. For example, hybrid approaches are classified as advanced and content-based and collaboration-based approaches as simple because hybrid approaches outperform the last two in terms of recommendation results coverage (Pazzani, 1999; Balabanovic & Shoham, 1997) and accuracy (Breese et al., 1998).

4.3 A Unified Framework for Improving Navigation

4.3.1 A Simplified Model for Personalising Web Directories

A personalised directory represents only the information (categories) in which users are interested. We present a simplified approach to personalise Web directories based on the use of explicit user profiling and content-based filtering techniques.

For the user profiling process, we decided to take simple factual data such as the websites a user visited in the past to learn their topics of interest. Such data can be obtained from their Web browsing history. Consider the experimental and psychological factors (i.e., intrusiveness, data privacy and trustworthiness) in the data collecting process, the construction process was set to *explicit*. That is, a user is asked to review their history list and choose the websites they want to use in their profile.

For the content filtering process, we decided to treat users *individually*. This means we only make category recommendations based on each user's profile instead of user group profiles. If the fundamental goal of this approach, which is to tailor a Web directory to present users interested categories based on their profiles, could be fulfilled, we may later expand this approach with collaborative filtering techniques.

4.3.2 A Redefined Search Model for Locating Categories

The OSM case study for identifying misfits in the Open Directory (Chapter 3, Section 3.2.6) suggested that the search engine of a Web directory should represent its use as a category locator rather than a general single-piece document finder like what a Web search engine does. This is because a Web directory is used to guide a user to locate the topic (category) they are interested and to direct them to explore it on the Web through the representative resources in the category. When categories are aligned hierarchically in a Web directory, the top-bottom browsing mechanism does not always work well as the user may show some understanding misfits on the structural arrangement of the directory. Therefore, the search engine of a Web directory should aim to help them find the topic directly no matter which super-ordinate categories it belongs to or how deeply it is "hidden". Since keywords have been artificially refined in Web directories to maintain the consistency and quality of the content (Chapter 3, Section 3.2.5.3), it is better to consider a search mechanism similar to a library catalogue that uses titles stored in the book shelves to locate these book shelves. In order to achieve this, we need understand how a search engine could help user navigate in a Web directory.

A generic Web directory normally organises websites in a general to specific order by using a hierarchical classification scheme. In other words, its top-level categories present the most general subjects while other level categories (i.e., ordinary categories and end categories) where website entries are collected, present more specific subjects. This structural arrangement not only helps a user specify their interest level by level but also determines that the main user's information search mode allowed on the directory should be browsing: a guided and semi-structured information search (Ellis, 1989). The advantage of browsing is that it supports both intentional (e.g., a user knows what he is looking for) and unintentional information request (e.g., a user does not know what he is looking for). On the other hand, Ellis (1989) points out keyword searching can only be done with intentional information request only. In other words, a user not only knows what they are looking for but also understands what kind of results they expect. In order to construct an objective query based on keywords, they must have good knowledge for their intention in terms of their expectation. Thus, if the user cannot establish such a connection between the information request, proposed results and keywords, the outcome of a search request will always be unsatisfying. To the search engine of a Web directory, it means that, first, a user must have an intention of interest. Second, the user must have some sort of expectation about the search results. For example, if they are looking for websites selling CDs and DVDs, in the lowest level, they need to know what a possible result could be. In this case, an online entertainment retailer. Finally, they need to know what kind of query can be used to reflect the first two prerequisites. Research on information seeking behavioural models states that querying is a searching stage after a user has “foreseen” some instances of results he is expecting (Bates, 1989; Ellis, 1989; Borgman, 1986, 1996; Ellis & Haugan, 1997; Kuhlthau, 1991,1993, 1994; Wilson, 1999 & 2000). Otherwise, they cannot formulate a query to describe their information needs completely. The ability of foreseeing expected results is normally obtained from a person's past knowledge or previous seeking activities in order to structure his information need and make it concrete. For example, a user has just read an article about automotive industries and they may have an interest for finding out about big multinational auto makers. If Volkswagon and Ford are mentioned in the article or the user knows them as famous automakers, they can compile an accurate query by stating VW or Ford as examples. In this case, the query could be “big multinational automakers + Ford or Volkswagon”. However, such queries seem not to be easily made through a directory's search engine as a Web directory often applies strict rules of language usage for identifying and categorising subjects as well as describing entries (websites). This maintains the authority and quality of a Web directory but also constrains the naturalness of keywords being used. Figure 4.5 shows a query of “big multinational automaker” to Open Directory which returns no results.

Figure 4.5 Search Results for "*big multinational automaker*" in the Open Directory

Retrieved March 2, 2008 from **http://search.dmoz.org/cgi-bin/search?search=big+multinational+automaker**

A generic Web directory like Yahoo! Directory or Open Directory is something akin to a huge reference library which aims to be useful to all users. Open Directory (2008) states that it normally only includes popular websites in categories for their high representative. Thus, we assume that, instead of using keywords, as long as a user knows something which can be seen as an exemplar of their expected results, he can use it to locate the corresponding categories. We call this as a "name-space" match mechanism. To some extent, it is similar to the library catalogue search.

4.3.3 The Architecture of Framework

The unified framework features an enhanced browsing model for generating personalised directory views based on user profiles and a redefined search model for locating categories based on user expectations of information need, as illustrated in Figure 4.6. Both sub-models share the use of a name-space content matching mechanism which matches a user's URLs to corresponding categories in the directory.

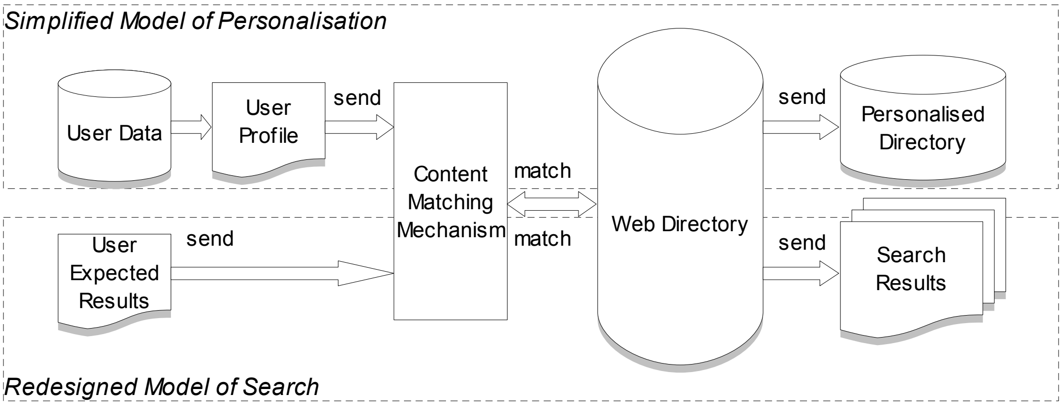


Figure 4.6 The Architecture of the Unified Framework

4.4 Implementation

4.4.1 Components Design

In order to implement the whole framework, we have designed and developed a number of components from emulating the environment to implementing the functionality of each model.

4.4.1.1 XML Parser

The Open Directory is an open-source project and it offers RDF dumps of its database with constant updates, which are available at http://rdf.dmoz.org/. For each set of RDF dumps, there are two types of raw RDF files.

Structure.rdf.u8.gz contains information about the category hierarchies. Figure 4.7 shows a short sample of the file where “<Topic/>” represents for a category, <d:Title/> is the title of the category, “<d:Description>” contains the HTML description of the category and each <narrow2/> or <narrow1/>indicates a child category. This file is mainly used to reproduce the whole representation of the Open Directory without entries.


```
<Topic r:id="Top/Arts">
        <catid>2</catid>
        <aolsearch>art</aolsearch>
        <dispname>Arts and Entertainment</dispname>
        <d:Title>Arts</d:Title>
        <d:Description><img src="http://dmoz.org/img/moz/mzcolor.gif" alt="Image from Mozilla
        museum: Mozilla as an Artist" width="128" height="120" align="right"> <p>The ODP <b>Arts</b>
        category contains English language sites about art, or "the use of skill and imagination in the
        creation of aesthetic objects, environments, or experiences that can be shared with others." This
        includes the "liberal arts," concerned with skill of expression in language, speech, and reasoning,
        and the "fine arts," concerned with affecting aesthetics directly, and especially affecting the
        sense of beauty. <small>(Quotes and paraphrases from <a
        href="http://www.britannica.com/">Britannica.com</a>)</small><p>Art is an abstract and
        subjective quality: It can be studied, but cannot be objectively measured, counted, weighed, or
        absolutely compared; it can only appeal to the viewer's or audience's personal
        senses.</d:Description>
        <altlang r:resource="Welsh:Top/World/Cymraeg/Celfyddydau" />
        <lastUpdate>2004-05-01 23:55:04</lastUpdate>
        <symbolic2 r:resource="Theatre:Top/Arts/Performing_Arts/Theatre" />
        <narrow2 r:resource="Top/Arts/Movies" />
        <editor r:resource="julianthurgood" />
</Topic>
<Alias r:id="Theatre:Top/Arts/Performing_Arts/Theatre">
        <d:Title>Theatre</d:Title>
        <Target r:resource="Top/Arts/Performing_Arts/Theatre" />
</Alias>
```


Figure 4.7 A Short Sample in the File “structure.rdf.u8.gz”

Retrieved August 25, 2007 from http://rdf.dmoz.org/rdf/structure.example.txt

Content.rdf.u8.gz contains links within each category. A short sample can be found in Figure 4.8 where “<Topic/>” represents for a category, “<catid/>” refers to the category's id and “<link/>” refers to an entity of the category. Note each <ExternalPage/> corresponds to each “<link/>” which contains detailed information about the entity such as “<d:Title/>”, “<d:Description/>” and “<topic/>”. This file can be used to reproduce a "pure" Open Directory without cross-references, language options and category descriptions.

```
<Topic r:id="Top/Arts/Movies/Titles/1/10_Rillington_Place">
        <catid>205108</catid>
        <link r:resource="http://us.imdb.com/title/tt0066730/" />
</Topic>
<ExternalPage about="http://us.imdb.com/title/tt0066730/">
        <d:Title>IMDb : 10 Rillington Place (1971)</d:Title>
        <d:Description>Full cast and crew for the film, and other information from the Internet Movie
                Database.</d:Description>
        <topic>Top/Arts/Movies/Titles/1/10_Rillington_Place</topic>
</ExternalPage>
```

Figure 4.8 A Short Sample in the File "content.rdf.u8.gz"

Retrieved August 25, 2007 from http://rdf.dmoz.org/rdf/content.example.txt

Since the core content matching mechanism is used to compare user inputs and user profiles with the URL entries in the directory, XML Parser (Figure 4.9) is a PERL script used for parsing "content.rdf.u8.gz" and extracting necessary nodes from the RDF file into a simplified XML file.

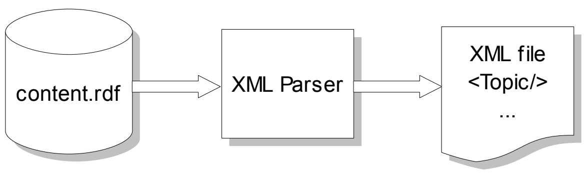


Figure 4.9 The XML Parser

A short sample of this XML file is shown in Figure 4.10. Compared to "content.rdf.u8.gz", the XML file only contains "<Topic/>" nodes where each "<Topic/>" node only has "<link/>" as its child nodes.

```
<Topic r:id="Top/Arts/Movies/Titles/1/10_Rillington_Place">
        <link r:resource="http://us.imdb.com/title/tt0066730/" />
</Topic>
```

Figure 4.10 A Short Sample of the XML Output for the File "content.rdf.u8.gz"

4.4.1.2 MySQL Database Importer

MySQL Database Importer (Figure 4.11) is another PERL script used for importing all nodes in the customised XML file into a MySQL database and indexing link data for search optimisation.

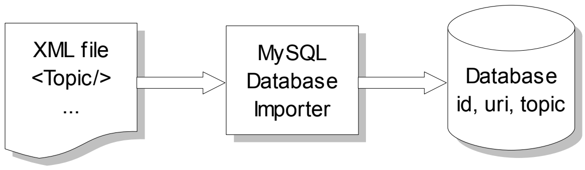


Figure 4.11 The Database Indexer

The database contains one table with three columns (“id”, “uri” and “topic”) where the “uri” column is indexed (Figure 4.12).

id (type: int(8))(KEY)	uri (type: TEXT)	topic (: TEXT)
1	http://us.imdb.com/title/tt0066730/	Top/Arts/Movies/Titles/1/10_Rillington_Place

Figure 4.12 The View of Database

4.4.1.3 User Profiling Agent

User Profiling Agent (Figure 4.13) is a Visual C++ application for extracting user profiles based on their browsing history.

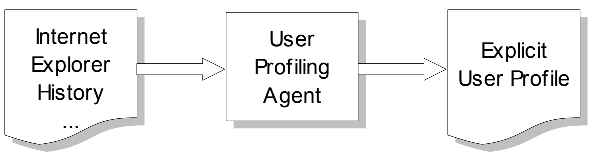


Figure 4.13 The Profiling Agent

In this implementation, Microsoft Internet Explorer 7 is selected as the default Web browser. Figure 4.14 shows a sample of user history accessed via some internal commands of the Internet Explorer.

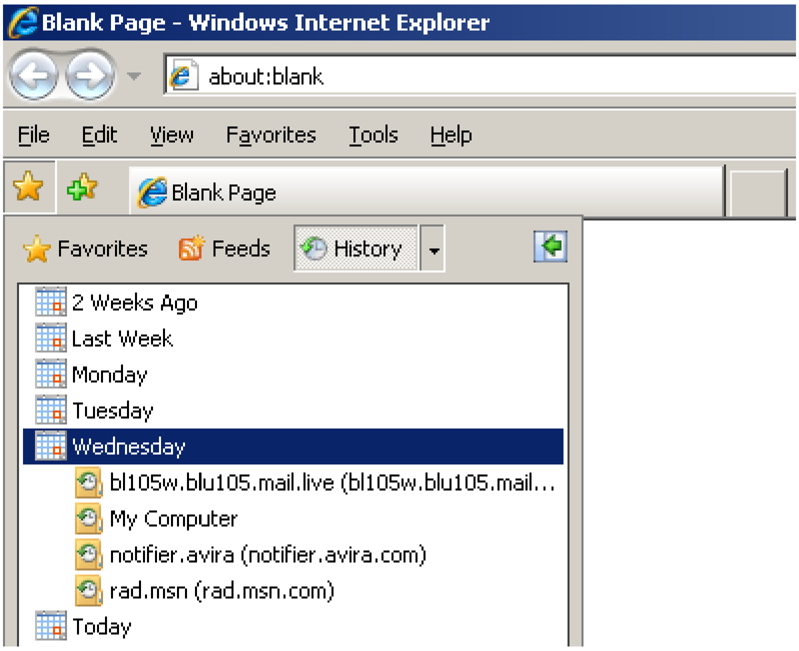


Figure 4.14 A Sample of User History in Microsoft Internet Explorer 7

The agent works like this: first, it reads user history in the Web browser and groups links based on their domains; then, it counts the total user visits of all sub-links in the same domain and re-orders the domains by total visits. We consider only the main domain (e.g., http://www.amazon.com/) as a valid URL instead of the whole link itself or a sub-domain of the link (e.g., http://www.amazon.com/ebooks/) as described in Chapter 1, Section 1.1 Footnote 1. The reason is that categories can be only named when they are knowledge domains or have become a popular phenomenon (e.g., Google as search engine culture) on the Web. However, the content contained by most Web links is too small and specific (e.g., a product page, or a news page) to be considered as a category of Web directories. A solution will be provided in the content matching mechanism in the next section for deciding whether the content of a Web link is representative-enough to match the subject of a category. Figure 4.15 shows a sample user profile. Since we decided to use explicit user profiles in case of user privacy, users are allowed to review their profiles and choose the domains they want to be "exposed" as their interest.

Domain	Counts
google.com	465
bbc.co.uk	65
dpreview.com	32
gizmodo.com	10

Figure 4.15 A Sample User Profile

4.4.1.4 Content Matching Agent

Content Matching Agent (Figure 4.16) is a PERL search script used to match user queries or user profiles to entries in the database with predefined match patterns and retrieve matched results. There are two search patterns defined, *exact match* and *expanded match*. Exact match is used to run an exact match between user input URLs and directory entries. This pattern is similar to Google's "I'm Feeling Lucky" option as it only returns categories containing exactly the same URL as user inputs. For example, when the user inputs "http://www.amazon.com/", category "Shopping: Entertainment" will be returned. However, expanded match performs a fuzzy search between user input and directory entries with the expansion to categories containing the URL appearing as a part of their entries. In the same example, "http://www.amazon.com/" will return categories like "Computers: E-Books: Readers" for "**http://www.amazon.com/**kindle/" and "Shopping: Publications: Digital" for "http://www.amazon.com/ebooks/". This is similar to Google's normal search option. Both match patterns utilise MySQL's default search options. Note only exact match is allowed for the personalisation.

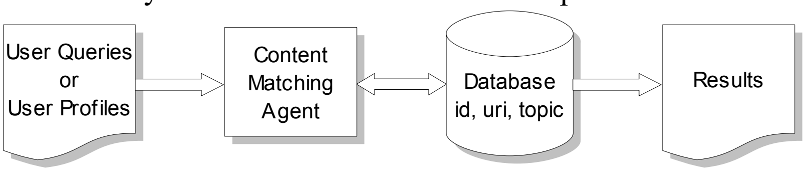


Figure 4.16 The Content Matching Agent

4.4.2 Hi-Fidelity Prototype

4.4.2.1 General Interface

The homepage of DMOZ was modified by adding a "Personalise!" button on the top right corner for reading and matching user profiles and a category locator search above of top-level categories.

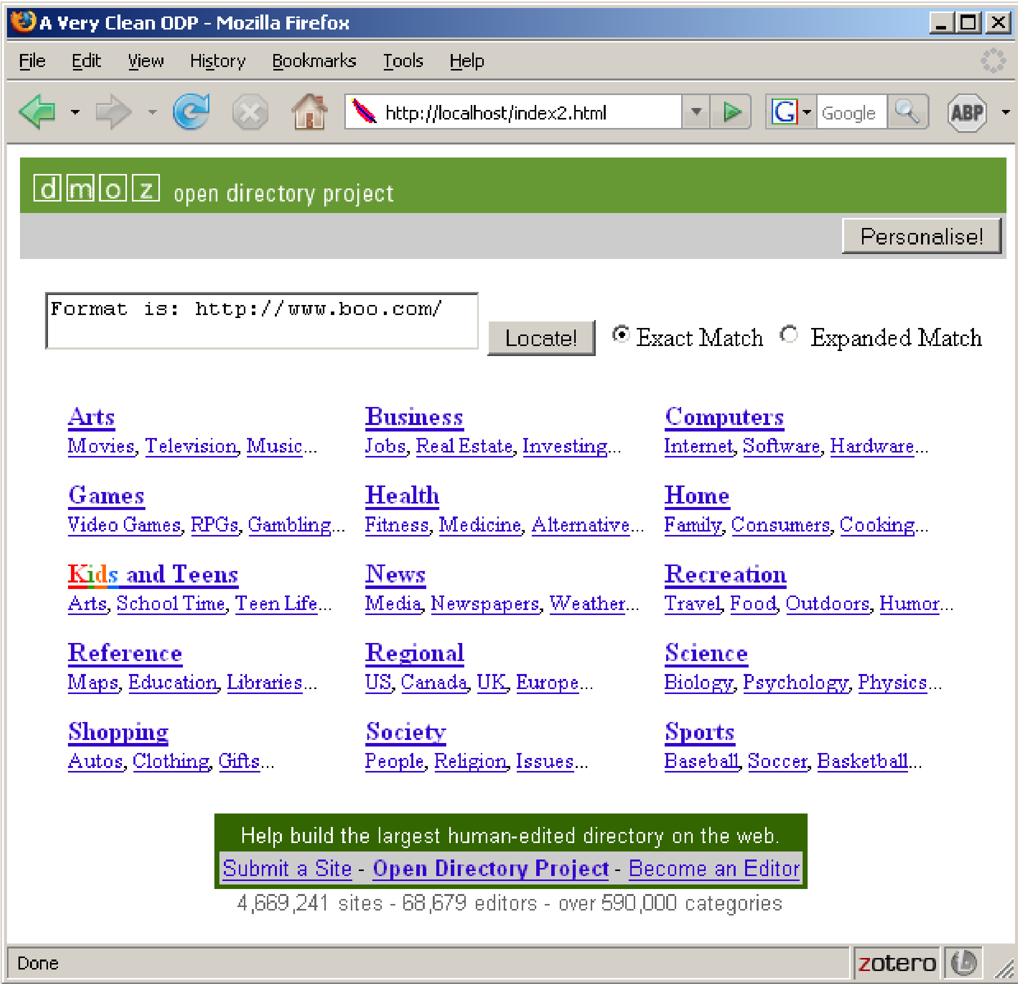
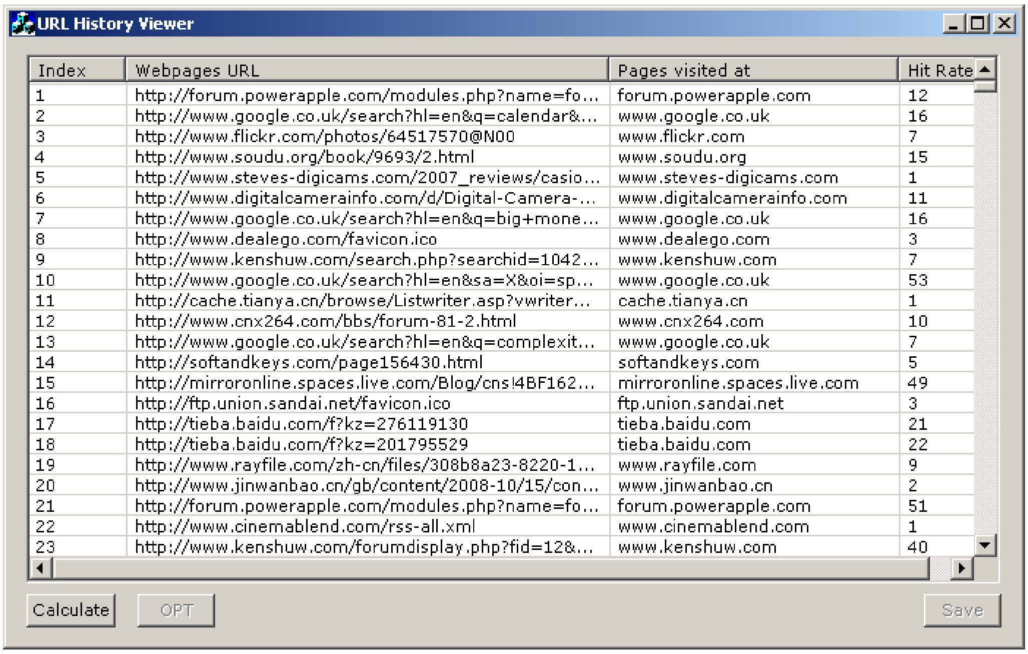


Figure 4.17 The Modified Homepage of the Open Directory

4.4.2.2 User Profile Generator

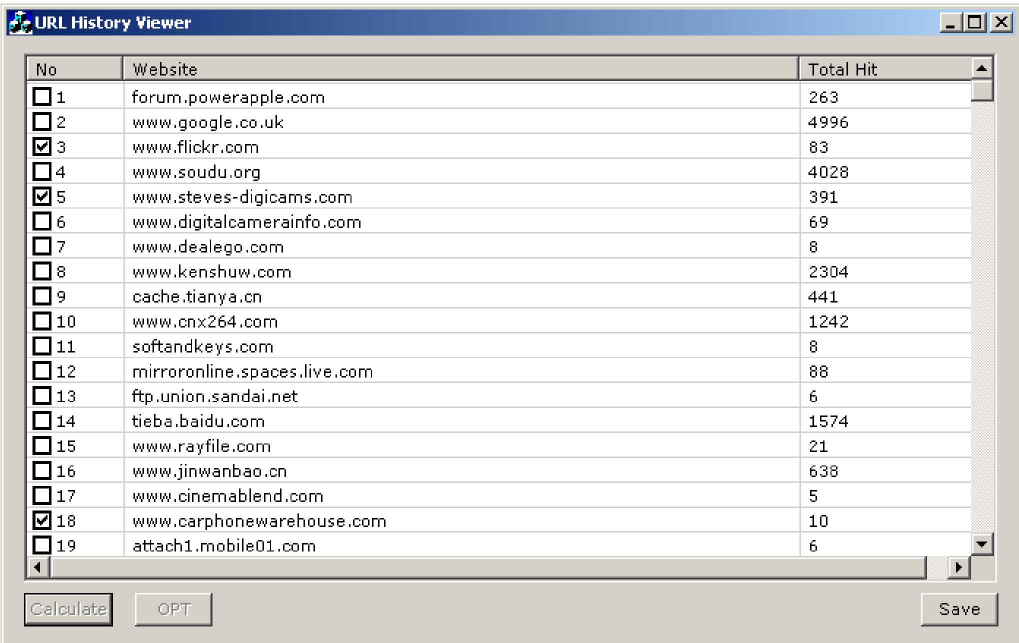
We implemented the user profile generator as “URL History Viewer” (Figure 4.18). First, the program will try to retrieve all user history in the Microsoft Internet Explorer and list them with “Index”, “Webpages URL”, “Pages visited at” and “Hit Rate”.



Index	Webpages URL	Pages visited at	Hit Rate
1	http://forum.powerapple.com/modules.php?name=fo...	forum.powerapple.com	12
2	http://www.google.co.uk/search?hl=en&q=calendar&...	www.google.co.uk	16
3	http://www.flickr.com/photos/64517570@N00	www.flickr.com	7
4	http://www.soudu.org/book/9693/2.html	www.soudu.org	15
5	http://www.steves-digicams.com/2007_reviews/casio...	www.steves-digicams.com	1
6	http://www.digitalcamerainfo.com/d/Digital-Camera-...	www.digitalcamerainfo.com	11
7	http://www.google.co.uk/search?hl=en&q=big+mone...	www.google.co.uk	16
8	http://www.dealego.com/favicon.ico	www.dealego.com	3
9	http://www.kenshuw.com/search.php?searchid=1042...	www.kenshuw.com	7
10	http://www.google.co.uk/search?hl=en&sa=X&oi=sp...	www.google.co.uk	53
11	http://cache.tianya.cn/browse/Listwriter.asp?vwriter...	cache.tianya.cn	1
12	http://www.cnx264.com/bbs/forum-81-2.html	www.cnx264.com	10
13	http://www.google.co.uk/search?hl=en&q=complexit...	www.google.co.uk	7
14	http://softandkeys.com/page156430.html	softandkeys.com	5
15	http://mirroronline.spaces.live.com/Blog/cns!4BF162...	mirroronline.spaces.live.com	49
16	http://ftp.union.sandai.net/favicon.ico	ftp.union.sandai.net	3
17	http://tieba.baidu.com/f?kz=276119130	tieba.baidu.com	21
18	http://tieba.baidu.com/f?kz=201795529	tieba.baidu.com	22
19	http://www.rayfile.com/zh-cn/files/308b8a23-8220-1...	www.rayfile.com	9
20	http://www.jinwanbao.cn/gb/content/2008-10/15/con...	www.jinwanbao.cn	2
21	http://forum.powerapple.com/modules.php?name=fo...	forum.powerapple.com	51
22	http://www.cinemablend.com/rss-all.xml	www.cinemablend.com	1
23	http://www.kenshuw.com/forumdisplay.php?fid=12&...	www.kenshuw.com	40

Figure 4.18 The Start Page of URL History Viewer

By clicking the "Calculate" button, it will sort webpages into their domains and add the total counts from each page (Figure 4.19).



No	Website	Total Hit
☐ 1	forum.powerapple.com	263
☐ 2	www.google.co.uk	4996
☑ 3	www.flickr.com	83
☐ 4	www.soudu.org	4028
☑ 5	www.steves-digicams.com	391
☐ 6	www.digitalcamerainfo.com	69
☐ 7	www.dealego.com	8
☐ 8	www.kenshuw.com	2304
☐ 9	cache.tianya.cn	441
☐ 10	www.cnx264.com	1242
☐ 11	softandkeys.com	8
☐ 12	mirroronline.spaces.live.com	88
☐ 13	ftp.union.sandai.net	6
☐ 14	tieba.baidu.com	1574
☐ 15	www.rayfile.com	21
☐ 16	www.jinwanbao.cn	638
☐ 17	www.cinemablend.com	5
☑ 18	www.carphonewarehouse.com	10
☐ 19	attach1.mobile01.com	6

Figure 4.19 The Sorting Page of URL History Viewer

Then the user can check the domain they want to add into their profiles and use "Save" button to save them into a CSV file (Figure 4.20).

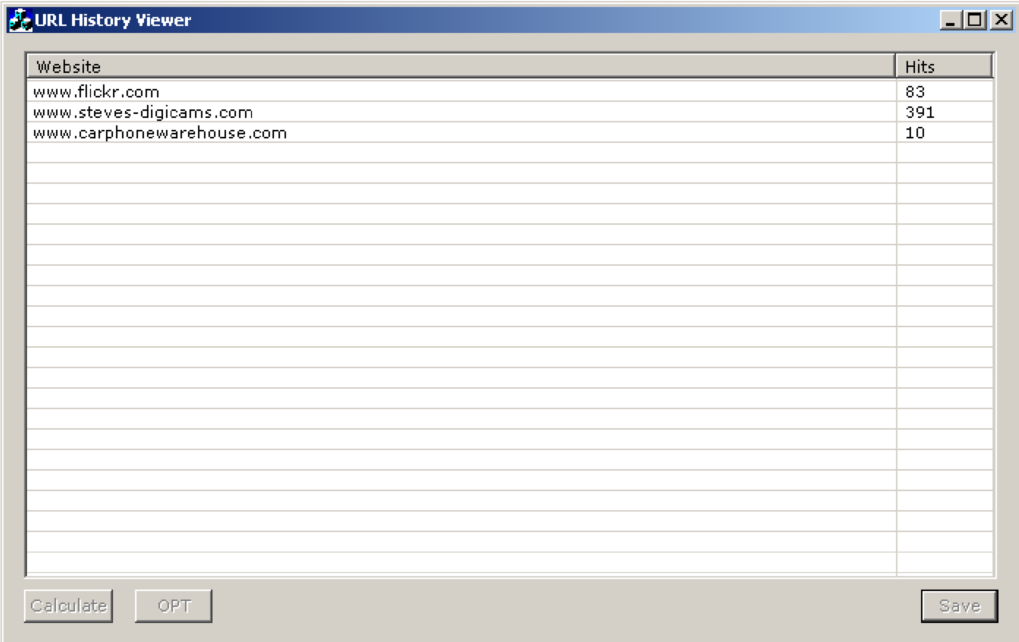
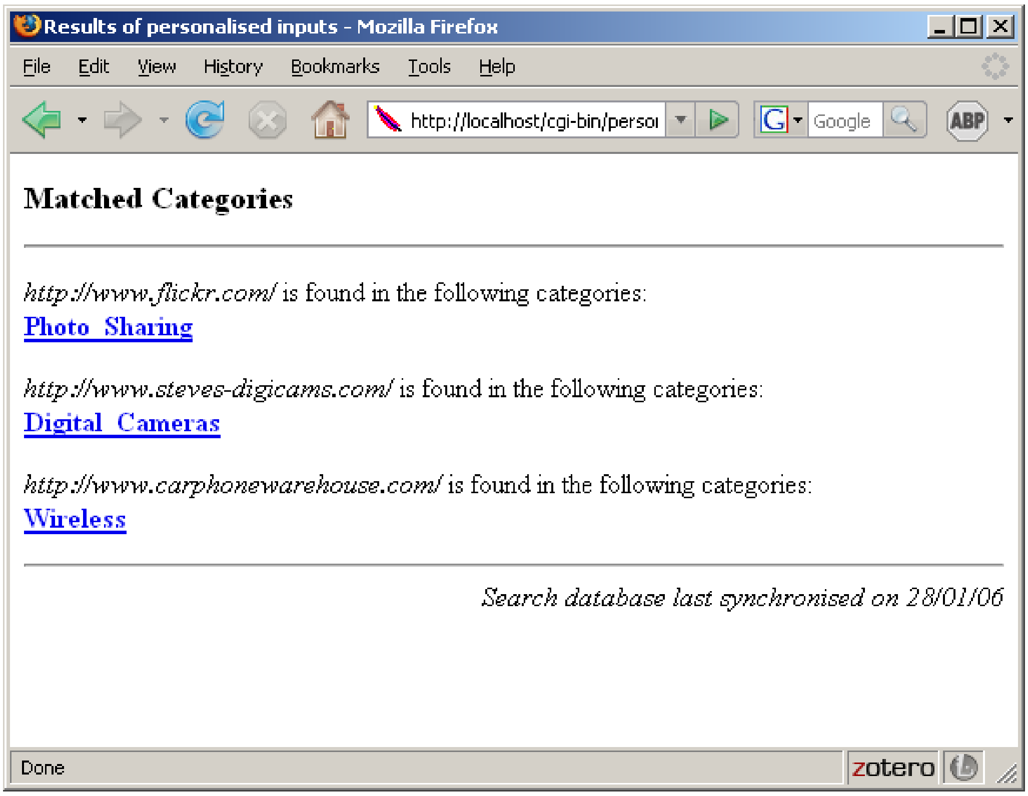


Figure 4.20 The Output Page of URL History Viewer

4.4.2.3 Content Match Results

Either searching through the category locator or using the personalisation based on user profiles (CSV files generated from URL History Viewer), a results page will look like below (Figure 4.21).



4.21 A Sample of Results

4.5 Summary

In this chapter, we first reviewed common techniques in terms of two main processes involved in Web personalisation, user profiling and content filtering. We then proposed a simplified model used for personalising Web directories in order to present directory content in a user-oriented view based on the user's interests. We also constructed a search model for locating categories based on the findings of the OSM study after studying typical user search behaviours on Web directories. Last, we combined our approach of the personalisation model and search model and illustrated the proposed architecture as a general framework for improving user navigation in the representation of Web directories. Moreover, we also outlined a number of agent developments in the prototype implementation. In the next chapter, we focus on an experimental design which was used for evaluating the implemented prototype to study how useful the framework can be.

Chapter 5 Experimental Design: A Comparative Usability Test

Usability refers to the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. (ISO 9241-11:1998)

5.1 Introduction

We have proposed a unified framework consisting of a redefined search model and an individual content-based personalisation model for improving the user experience of navigation in Web directories. We also implemented this framework as D-Search for the redefined search model and D-Persona for the personalisation model on the basis of the Open Directory. In this chapter, we plan to conduct a comparative usability evaluation for studying how “useful” the framework is compared to not only the original Open Directory but also Google. This is due to the fact that search engines have replaced Web directories in the role of helping people guide and locate new websites. The comparative study is summarised in the Table 5.1 where detailed designs are discussed in the specified sections.

<table>
<tr><th>Designed Tasks</th><th>Systems involved</th><th>Executing people</th><th>Sections</th></tr>
<tr><td>User grouping</td><td>Search engines (generic)</td><td rowspan="3">Participants & observer</td><td>5.5.1 & 5.5.2</td></tr>
<tr><td>Background survey</td><td>Web directories (generic) & search engines (generic)</td><td>5.5.2</td></tr>
<tr><td>Level test 1 and 2</td><td>The Open Directory & Google</td><td>5.5.2</td></tr>
<tr><td>Simple tasks for D-Search</td><td rowspan="2">D-Search, the Open Directory & Google</td><td rowspan="2">Participants</td><td rowspan="2">5.6.3</td></tr>
<tr><td>Complex tasks for D-Search</td></tr>
<tr><td>Task data collection for D-Search tasks</td><td></td><td>Observer</td><td>5.6.1 & 5.6.2</td></tr>
</table>

Open task for D-Persona	The Open Directory & D-Persona	Participants	5.7.2
Task data collection for D-Persona	D-Search, D-Persona, the Open Directory & Google	Observer	5.7.1
User feedback ratings	D-Search, D-Persona, the Open Directory & Google	Participants	5.6.2 & 5.7.1
User open-text feedbacks	D-Search & D-Persona		

Table 5.1 An Overview of the Task Design

5.2 Usability: The Definition of "Usefulness"

The formal term used for describing the usefulness of a design is usability, which denotes how easy and quick people can use a product to accomplish their tasks (Dumas & Redish, 1993). The term usability also refers to the methods of measuring usability for improving the ease-of-use of a design, which are composed of five quality metrics (Nielsen, 1994):

- Learnability: How easy is it for users to accomplish basic tasks the first time they encounter the design?
- Efficiency: Once users have learned the design, how quickly can they perform tasks?
- Memorability: When users return to the design after a period of not using it, how easily can they re-establish proficiency?
- Errors: How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
- Satisfaction: How pleasant is it to use the design?

Web usability is an application of usability in domains where Web browsing can be considered as a general metaphor for constructing user interface. Usability.gov (2009) points out there are three measurements need to be considered when conducting usability testing for websites in compliance with the general usability measurement.

- Effectiveness: Can a user successfully use a Web site to find information and accomplish tasks?

- Efficiency: Can a user quickly accomplish tasks with ease?
- Satisfaction: How much does a user enjoy using the Web site?

Data captured during user testing are in two types, performance data (what actually happened) and preference data (what participants thought). Performance data are normally used for measuring efficiency and effectiveness while preference data are used for measuring user satisfaction. Nielsen (2003) recommends choosing representative users and representative tasks with sufficient user observation for maintaining data accuracy and validity for the measurement, where:

- Representative users are typical users that would use the interface such as customers for an e-commerce site;
- Representative tasks are typical tasks that users would perform on an interface such as placing an order on an e-commerce site;
- User observation covers topics including what the users do, where they succeed, and where they have difficulties with the user interface.

5.3 The Purpose of User Testing

The primary purpose of our user test is to assess the usability of the unified framework as a user-centred solution for improving the user-system misfits in the user navigation of Web directories. The secondary purpose was to identify whether the unified framework can support the purpose of using Web directories (i.e., guiding users to locate websites). In addition, user satisfaction and preference with the unified framework were also investigated in comparison to the original Web directories and search engines for understanding the possibility of introducing this solution into Web directories.

5.4 Hypotheses

The unified framework consists of two sub-models with different objectives, which includes a redefined search model for enhancing users' search experience and an individual content-based personalisation model for improving users' general browsing experience. Although they share the same content-matching mechanism, they are separate models addressing different usability misfits. Therefore, hypotheses for the experiment will be explained based on D-Search and D-Persona respectively. In addition to the hypotheses designed for verifying our primary purpose and secondary purposes in terms of the Web usability measurements (Section 5.2), we also added a "helpfulness" hypothesis test for each sub-model in order to justify the user preference of using the unified framework compared to the original support of browsing and search.

5.4.1 Hypotheses for D-Search

The OSM usability inspection study in Chapter 3 suggests that a Web directory should provide a search facility for locating its categories more effectively rather than using a simple Web search engine to find websites. Following this suggestion, we redefined the user search model of Web directories and implemented a category locator which allows users to submit websites they know as exemplars (expected results) of their interests for locating relevant categories containing them. The theory is based on two points. First, the websites of a category in a Web directory, as the entities of a class in a hierarchical classification, are the most ideal descriptors and representatives for the category because they inherit all attributes defined by the category and present them as real exemplars. Second, since a Web directory only collect representative websites where most of them are popular and reputable, such entities are generally not difficult to be recalled from users knowledge compared to using self-complied keywords to describe the interests.

Thus, for directory-featured information needs, our hypotheses are described below.

> ***H1:*** *For the efficiency, with D-Search, a user will complete search tasks quicker than using the original Open Directory and Google.*
>
> ***H2:*** *For the effectiveness, with D-Search, a user will complete search tasks more successfully than using the original Open Directory and Google.*
>
> ***H3:*** *For the user satisfaction, with D-Search, a user will be more satisfied in completing tasks with the Open Directory than before.*
>
> ***H4:*** *For the helpfulness, a user will find using D-Search to search the content of the Open Directory is more helpful than its original search engine or Google in terms of their intentions.*

5.4.2 Hypotheses for D-Persona

The basic idea of using personalisation as an approach for improving the user experience in Web directories is that user could always encounter with understanding difficulties during their navigation no matter how well these directories are organised and represented. This is because generic Web directories focus on general use only which makes them difficult to reflect individual demands. So when a user needs to make extra effort in understanding and distinguishing information content in Web directories, some understanding problems will appear and their navigation will be affected. Thus, if a Web directory could offer some kind of representation that is based on each individual's need, these issues which cause problems in user navigation would be minimised. Thus, our hypotheses are:

> ***H1:*** *For the efficiency, with D-Persona, a user will find his interested content quicker than with the original Open Directory.*
>
> ***H2:*** *For the effectiveness, with D-Persona, a user will find his interested content more successfully than with the original Open Directory.*
>
> ***H3:*** *For the satisfaction, with D-Persona, a user will be more satisfied with the*

Open Directory.

H4: *For the helpfulness, a user will find D-Persona could improve their judgement in the helpfulness of the Open Directory in terms of its usage.*

5.5 Deciding Test Participants

5.5.1 Defining Representative Users

Defining test participants is a key process of user testing along with setting up measurement for user observation and deciding representative tasks. Nielsen (2003) recommends choosing representative users as test participants in order to maintain data accuracy and validity for the measurement. For our comparative user study, representative users could be any Web users with the need for locating websites. However, due to the fact that most Web users use search engines instead of Web directories, we defined that the representative users for this test had to be search engine users at least. We then recruited them from Queen Mary, University of London where most came from the Department of Computer Science for the convenience.

5.5.2 Grouping Test Users

Nielsen (2000) reminds that Web usability has traditionally been focused on increasing ease of learning for the novice users and he stated this should continue to be main goal for any Web usability study. That is, participants of this experiment should ideally be novice Web directory users. However, we decided to have two user groups based on their search engine expertise. This is mainly because our redefined search engine is essentially a search facility which is similar to a normal search engine in supporting user search strategies (e,g, formulating an interest with short descriptions and browsing the search results etc.) although the search model behind it is redefined exclusively for searching in Web directories. In this respect, an experienced search engine user will receive more benefits and thus perform differently from a search engine user with little experience when they use the redefined search engine. As Web directories are not as popular as search engines, it is very likely that two search engine users with different level of search experience are classified as novice Web directory users. If this happens, the two users would perform differently and affect our understanding of the results. Therefore, we prepared a user questionnaire (Appendix 1.1) to understand the users' Web behaviour and background knowledge of the Web. We also designed a user search task to identify their level of search experience (Appendix 1.1) based on their performance and then classified them into two levels of user groups: the novice user group and expert user group. Moreover, a user task for locating a category in the Web directory was introduced for understanding their knowledge of classifications.

5.5.3 Choosing The Number of Users to Test

For qualitative analysis, Nielsen (2000), Nielsen & Landauer (1993), Virzi (1992) and Lewis (1994, 2006) claim that, based on mathematical models and empirical evidence for the models, using small sample size like five participants could detect most usability problems (Virzi: approx. 80% and Nielsen: 84%) in a product cost-effectively. However, Woolrych & Cockton (2001) criticise that small user sets are not reliable as there is no way to determine that any set of five tests matched those percentages, or which particular problems were revealed or missed. Faulkner (2003) also argues that the assumption depends on the independence of the problems encountered – that is, that encountering one of them will not affect the probability of encountering any other problem. In one study (Spool & Schroeder; 2001), the first five users revealed only 35%

usability problems and both the 13th and 15th users revealed at least one new and severe usability problem. In another study with 18 users (Perfetti & Landesman; 2002), each new user, including those in test sessions 6 – 18, found “more than five new obstacles”. These experimental results suggest us deciding the sample size of qualitative studies based on the number of participants required for quantitative studies. Nielsen (2006) recommends testing with 20 participants for hitting in ±19% confidence interval to the margin of error in practice when collecting quantitative usability metrics. In addition, Faulkner's usability study (2003) with 60 users sampled from three levels of user experience reveals that each set of randomly selected 20 users could find 95% usability problems. Based on these guidelines, we decided to recruit 24 – 30 participants to run the test.

5.6 Measuring D-Search

D-Search is an implemented search engine which reflects the redefined search model and offers users the ability to locate categories based on their expected results (website exemplars). We decided to capture both performance data and preference data for measuring the effectiveness, efficiency and user satisfaction of D-Search in a comparative study with the original search method of Open Directory and Google in specific search tasks.

5.6.1 Capturing Performance Data

5.6.1.1 Task Completion Time

Task completion time or time on the task is a performance metric for measuring the efficiency of a system like how quickly users are able to accomplish tasks on the system. A user's task completion time is calculated from the start of a task to the end of the task. The average task completion time on a system is calculated by taking the average time of a group of users for using the system to complete the same tasks. Let $\boldsymbol{T_{System}}$ represent ***the average task completion time*** of ***n*** users on a system and $\boldsymbol{t_{System}}$ represent the task completion time of a user, then $\boldsymbol{T_{System}}$ is:

$$T_{System} = \overline{\sum_{i=1}^{n} t_{system}} (n \geq 1) \quad [1]$$

Since the calculation is based on the time a user spent on a task without considering whether the task is successful or not, the user could spend very short time on the task if they lack confidence to complete it and then declare a failure. In this case, an observer will judge whether it is an early announcement made by insufficient user efforts. If it is, the user will be asked to continue until their efforts are considered to be enough to make the decision.

5.6.1.2 Normalised Success Rate

Success rate is the simplest performance metric used for measuring users' ability to complete tasks on a system in terms of effectiveness. Nielsen (2001) describes a simplified scoring methodology for measuring success rate by classifying a task as success (1 credit), failure (0 credit) and partial success (where he recommends to give 0.5 credit in practical). For example, three users were asked to perform the same task and their results were *Success*, *Partial Success* and *Failure* respectively. Let Success be one credit, Failure be zero credit and Partial Success be 0.5 credit, the success rate is calculated by taking the average task credit of all three users, which is (1 + 0.5 + 0)/3 = 50%. This method gives a general perspective of how a system supports users and how much improvement is needed to make the system really work. However, there is no firm rule for assigning credit to partial success as the definition of partial success is varied from one test to another. If the wrong credit is assigned to a partial success, the accuracy of overall success rate will be highly affected, especially for a comparative user study where several systems are assessed with the same tasks. Thus, instead of defining a partial success and giving a credit to it, we suggested considering only success and failure of a task but using the success rate of user sessions of one user to normalise the task success rate. For search tasks, a user session could be considered as a query session that a user submits a query to the system and examines the returned results. Then the normalised success rate of a system can be calculated by using the following formula,

$$\overline{S_{NSR}=\sum_{i=1}^{n}\sum_{j=1}^{m} t_{SR}\times u_{SR}}\left(S_{SR}, t_{SR}, u_{SR}\subset 0,1 ; n, m\geq 1\right) \quad [2]$$

where S_{NSR} is the normalised success rate of a system, t_{SR} is the success rate of all users for one task, u_{SR} is the success rate of their user sessions of the task, n is the number of users and m is the number of tasks a user has performed.

5.6.1.3 User Pathway

In addition to task completion time and success rate, we also decided to introduce pathway analysis for conducting detailed measurement of each system. This can be done by studying certain factors of user search processes like the number of queries user used for completing a task on each system.

5.6.2 Capturing Preference Data

Collecting preference data for measuring user satisfaction on D-Search is achieved by studying user comments and preference ratings. Since the user tasks designed for D-Search are comparative search tasks, we prepared a selection of rating questions for analysing user preference of D-Search in comparison of the original Open Directory search and Google in addition to standalone ratings (Appendix 1.2).

5.6.3 Designing Search Tasks

The search tasks used for collecting performance data must be the typical tasks would perform on a system. In other words, these tasks should present the typical use of the system and/or they must be commonly performed by most of the representative users of the system. This can be done by determining task goals, choosing task topics and defining task complexity.

5.6.3.1 Determining Task Goals

For a comparative user testing, it needs to take into account whether the task is justified for balancing the strength of different testing systems so as to make them actually comparative. We decided to set the task goals to *finding a number of websites on specific topics* where the number of websites depends on the maximum possible number of websites acquired by the Open Directory and the topics depends on the extent of topics supported by the Open Directory. This is due to two reasons. First, although Web directories and search engines support the same way in searching information, Web directories are more limited than search engines in terms of their information coverage (Chapter 1, Section 1.2.2). Second, although both of them allow topic-based search, Web directories have more limited topic extent than search engines due to their establish purposes and classification schemes (Chapter 1, Section 1.2.3). Thus, we decided to use relatively broad search topics with website-level search goals.

5.6.3.2 Choosing Task Topics

We decided to use popular search interests as task topics for avoiding to collect unnecessary user data on the systems caused by the lack of understanding or misunderstanding of task topics. Thus, the following topics were selected based on an annual search interests report of Google.

1. Online shopping (media, consumer electronics, fashions etc.)
2. Online services (banking, broadband, insurance etc.) or utilities transfer.
3. Online booking (Entertainment, Holiday, flight etc.).

5.6.3.3 Task Complexity

We also need to consider task complexity in addition to choosing task topics and goals in defining representative tasks as the task performance relies squarely on user's understanding of a task (or problem) in terms of the intention of its information needs (Belkin et al., 1982; Ingwersen, 1992, Robinson, 2001). This view is supported by many studies in different fields (Locke et al., 1981; Wood et al., 1987; March & Simon, 1967; Van de Ven & Ferry, 1980; Culnan, 1983; Hart & Rice, 1991; Tiamiyu, 1992) where the relationships of various types of tasks and information needs have been widely investigated (Brittain, 1971, 1975; Dervin & Nilan, 1986; Tushman, 1978). Referring to task categorisation introduced by Byström & Järvelin (1995), two types of tasks, which are called *simple tasks* and *complex tasks*, were defined to present different level of search complexity. Generally speaking, they are different in the following aspects.

(1) The complexity of a task goal. A simple task normally has an easy-to-achieve goal while a complex task has a goal which is required to divide into sub-goals to achieve.

(2) The number of work sessions required in performing a task. A simple search task needs fewer queries and less information differentiation processes compared with a complex task.

(3) The number of different types of user actions required in performing a task. For instance, a simple task requires less query optimisation, query reset and results differentiation than a complex task.

(4) The required expertise for problem-solving. For example, experienced users can easily handle both simple tasks and complex tasks while novice users may produce significant varied time differences in performing simple tasks and complex tasks.

Detailed tasks are presented in Figure 5.1. . Due to the fact that the search results in a Web directory can only be a number of websites/pages presenting the relevant topic represented by the search query, we added the requirement of extra user effort to make tasks more complex and to reflect their difference from simple tasks in the above four aspects. That is, a simple task may have a goal to find out a number of websites having a similar topic but a complex task would require a user to obtain more information based on the websites they find. Taking the first complex task as an example, the task goal is to find out the cheapest possible retail price of a 512MB memory card. Users need to understand what the task goal requires and then take appropriate users actions and search strategies for problem-solving. First, the task goal indicates increasing number of work sessions as users need to compare the product price in a number of websites. Second, it requires advanced problem-solving skills as they need to choose a suitable starting point – this product is being sold in various types of websites like Argos, Tesco, Currys, Jessops or Play.com but which one has advantages over others? Ideally, users were required for perform at least one pair of tasks (i.e., a simple task and a complex task) with each system.

Simple tasks	Can you find out 20 online mobile phone retailer? Note a network carrier like Orange, O2, 3, Vodafone or Virgin Mobile cannot be considered as a mobile retailer.
	Can you find out 20 utility suppliers? Hint: utility suppliers mean gas and/or electricity suppliers.
	Can you find out 20 ISPs. Hint: ISP stands for Internet Service Provider, they are commonly known as broadband companies offering broadband packages for home/business users.
Complex tasks	Can you find out the possibly cheapest retail price of a Sandisk 512MB Memory Stick Pro (not Memory Stick Duo or Memory Stick) on the Web. Hint: in order to make sure the price is as low as possible, you need to compare the price you find with at least 4 other websites. Note a comparison website will count for one website only.
	Can you find out the possibly cheapest price of a 4-slot (4-slice) toaster (any brand) on the Web. Hint: same as above. Note: same as above.
	Can you find out the possibly cheapest return ticket from London to Paris, which departs on 18 April and returns on 20 April. Hint: same as above. Note: same as above.

Figure 5.1 The Pool of User Tasks

5.6.4 Recording Data

Relevant user test data were hand-recorded by an observer using a prepared recording sheet during the test (Appendix 1.3).

5.7 Measuring D-Persona

D-Persona is an implementation of the individual content-based personalisation model which aims for providing tailored content access experience for an individual user based on the analysis of their own histories or past activities. This indicates that the difficulty of comparing users' personalised results. Thus, we decided to capture preference data (e.g., how a user feels about personalisation) for measuring D-Persona rather than capturing performance data (e.g., how the personalisation performs in some specific search tasks).

5.7.1 Designing The Questionnaire

The preference data we decided to capture were subjective satisfaction (e.g., do users enjoy using the system) and user comments (e.g., are they confused by the system). A one-to-six user satisfaction rating system was used for collecting user satisfaction and an open question was asked for collecting qualitative user comments respectively (Appendix 1.2). We also added rating questions for understanding how users think of the importance of personalisation in terms of navigating in Web directories. Note questions comparing D-Persona to other similar services offered by existing Web directories were not added due to the fact that this function has not been seen in mainstream Web directories[16].

5.7.2 Designing The Open Task

Each participant was instructed to create their profiles based on their past browsing history and then to submit their profiles into a clone of the Open Directory for generating their personalised directories. They were also requested to run D-Persona as many times as necessary until they had gained a full understanding of D-Persona.

5.8 Suggestion from The Pilot Study

We scheduled a pilot study with one novice user and one expert user to check the timings and logistics of tasks. The study suggested reducing the amount of tasks performed by each user. Thus, we changed the initial plan that a user must perform one pair of tasks on each system to the actual plan that a user should perform one pair of task on two selected systems. This will make each user group perform 8 tasks on each comparative platform for each type of task in total. For example, the novice user group will perform 8 simple tasks on Google and 8 complex tasks on Google and then the same number of tasks on the Open Directory and D-Search each. Suppose there are 3 users in the novice user group. User 1 performed 1 pair of tasks on system 1 and system 2), user 2 performed 1 pair of tasks on system 2 and system 3 and user 3 performed 1 pair of tasks on system 1 and system 3. In this way, each system is used for 2 pairs of tasks (2 simple tasks and 2 complex tasks). Following the same calculation, for a group with 12 users, 8 pairs of tasks are performed in each system.

16 Representative Web directories include WWW Virtual Library, Best of the Web Directory, Starting Point Directory, JoeAnt.com, Ansearch Directory, Yahoo! Directory, Dmoz (The Open Directory Project) and directories powered by DMOZ such as Google Directory, Lycos Directory, AOL Yellowpage etc.

5.9 Summary

We have explained the whole design of our comparative user study in this chapter. Table 5.2 shows a summary of the experiment where the user tasks, estimated time and data collection methods for each user are explained. Here the estimated time (135 minutes) of all user tasks a user need to perform was calculated based on the users' performance in the pilot study when they were asked to complete tasks successfully. On some occasions, a user tended to reject a task if it looked complex after or even without a few attempts. If it was the case, the user was encouraged to try their best regardless of the time limits.

	User tasks		**Estimated time**	**Data gathering methods**
Session 1	User background research	General Interview	10 minutes	Questionnaire
		User test 1	15 minutes	Recording form (observation)
		User test 2	15 minutes	
	User tasks	Simple task 1 (D-Search)	45 minutes	Recording form (observation)
		Simple task 2 (D-Search)		
Session 2	User tasks	Complex task 1 (D-Search)	75 minutes	Recording form (observation)
		Complex task 2 (D-Search)		
		Open task (D-Persona)	15 minutes	Questionnaire
	User feedback	User ratings (D-Search & D-Persona)	15 minutes	Questionnaire
		Open text questions (D-Search & D-Persona)		

Table 5.2 A Summary of the Experimental User Study

Chapter 6 Results & Discussions

6.1 Introduction

In this chapter, we report the user results from the comparative user study where the results for D-Search and D-Persona are given separately based on the different measurements used. We then discuss these results in terms of the experimental hypotheses followed by a summary. Table 6.1 shows an overview of the arrangement.

	Results	Sections
Participators' background	Users' system expertise	6.2
	Users' task completion	
Results of D-Search	Task completion time	6.3.1
	Normalised system success rate	6.3.2
	User pathway data	6.3.3
	User feedback ratings	6.3.4
	Subjective Feedbacks	6.3.5
Results of D-Persona	User feedback ratings	6.4.1
	Subjective feedbacks	6.4.2
Discussions		6.3
Summary		6.5

Table 6.1 An Overview of the Chapter

6.2 Participators' Background

We recruited twenty-four 3rd year students as our testing participants and halved them into two equal sized user groups (the novice user group and expert user group) based on the results of their Web usage survey and search expertise tests. We also paid each student 20 pounds for their cooperation and time in the experiment.

The grouping was mainly determined by the results of users' search expertise test where they were required to find out the source of a research document referred in a recent (by the time of running this experiment) BBC programme called "Little Kinsey" on the top of an insight of their Web usage survey. From the Web usage survey, we had a general idea that expert users normally spent a longer time on the Internet every day and had a generally broader interests of topics and activities than novice users. The results of the two expertise tests are shown in Table 6.2 where it clearly shows that the expert user group performed search expertise tasks more quickly and successfully than the novice user group (251.31 seconds vs 308.75 seconds; 83% vs 50%). During the observation, we found that most users in the two groups were able to find the actual BBC link within a few searches at the beginning. However, the difference between the two groups is that most novice users who failed the task were unable to retrieve useful information from the link to form new queries to specify the search whereas only two expert users were unable to make progress from the BBC link. We also saw this experience-orientated difference in completing directory expertise test although the two groups reported the same success rate.

Group	Search Expertise Test		Directory Expertise Test	
	Completion time (s)	Success rate (%)	Completion time (s)	Success rate (%)
Novice	308.75	50.00%	122.16	83.00%
Expert	251.31	83.00%	49.11	83.00%

Table 6.2 User Expertise Test Results

The background survey also showed that all participants described themselves as regular search engine users and only 41.67% users from the expert group said they had used general Web directories like the Open Directory occasionally (Table 6.3).

Group	Size	Search Engines			Web Directories		
		Size (Percentage: %)			Size (Percentage: %)		
		Regular user	Occasionally used	Never used	Regular user	Occasionally used	Never used
Novice	12	12 (100%)	0 (nil)	0 (nil)	0 (nil)	0 (nil)	0 (nil)
Expert	12	12 (100%)	0 (nil)	0 (nil)	0 (nil)	5 (41.67%)	0 (nil)

Table 6.3 User Usage of Google and The Open Directory

Further, no participants reported frequent difficulties on search engines while 60% of directory users in the expert group reported such difficulties (Table 6.4). 41.67% in the novice group and 58.33% in the expert group did not have any difficulties when using search engines compared to only 20% of directory users having the same feedback.

58.33% of novice users and 41.67% of expert users said they had experienced difficulties occasionally with search engines compared to 20% with Web directories. In summary, results of Table 6.3 and 6.4 reflect the fact that search engines are much more popular used by normal Web users for online navigation than Web directories in these days.

Group	Size	Search Engines			Web Directories		
		Size (Percentage: %)			Size (Percentage: %)		
		Often	Occasionally	Never	Often	Occasionally	Never
Novice	12	0 (nil)	7 (58.33%)	5 (41.67%)	n/a	n/a	n/a
Expert	12	0 (nil)	5 (41.67%)	7 (58.33%)	3 (60%)	1 (20%)	1 (20%)

Table 6.4 User Encountered Difficulties with Google and The Open Directory

We also tested our participants' conceptual classification mapping to the Open Directory and found only 8.33% of novice users and 50% of expert users derived the same mappings for classifying Amazon.com[17] (Table 6.5). In this test, participants were asked to define the website's topic and specify a category containing it in the Open Directory. If some of them had not visited Amazon before, they were also asked to spend some time on it for deriving their understanding. Our results show that most users failed in continuing browsing from “Shopping: Books@” in the Open Directory as they expected subcategories like “Book Retailers” or “Book Stores” but the directory only has subcategories in terms of topics. For example, Arts, Fiction, Science etc. This finding indicated that the classification difference is a major misfit between users and Web directories to cause users' navigational difficulties. In addition, the better results reported by the expert group also reflected the suitability of grouping based on their search engine experience as most users did not have any experiences with Web directories.

Group	Size	Mapping Amazon's category in the Open Directory	
		Correct	Wrong
Novice	12	1 (8.33%)	11 (91.67%)
Expert	12	6 (50%)	6 (50%)

Table 6.5 Users' Conceptual Classification Mapping

The actual user task completion statistics on each comparative system are summarised in Table 6.6. As we explained in the pilot study findings in last chapter, the number of

17 In the Open Directory, Amazon.com is classified into 3 different subcategories of the top level category “Shopping”: “Major Retailers (under General Merchandise)”, “A (under Publications: Books: General Interest)” and “Entertainment”.

tasks performed on each system was eight instead of twelve.

Group	Size	Number of Tasks Completed						
		D-Search		Google		Open Directory		D-Persona
		Simple	Complex	Simple	Complex	Simple	Complex	Open Task
Novice	12	8	8	8	8	8	8	12
Expert	12	8	8	8	8	8	8	12

Table 6.6 Number of Tasks Completed by User Groups

The detailed results of task performance and user feedback are reported and then discussed in the following part. Since D-Search and D-Persona were tested with different user tasks and aims, their results are reported separately.

6.2 Results of D-Search

D-Search is an implementation of the redefined search model in our unified framework which aims to reflect users' true need of directory searching. The results of D-Search are reported into five measurements. They are, *Task Completion Time* for measuring the efficiency, *Normalised Success Rate* for measuring the effectiveness, *User Pathway Data* for detailed analysis, *User Feedback Ratings* and *Subjective Feedback* for measuring user satisfaction. Results in the first three measurements are reported in two views: group based analysis – presenting results within the same group and system based analysis – presenting results on the same system where different statistical analysis are used.

6.2.1 Task Completion Time

Task completion time measures the efficiency of a system in performing tasks by calculating the average time users take for completing designated tasks. The tables of results are generated by using SPSS where abbreviations are used to refer the three systems. For example, “OD” stands for “The Open Directory”, “G” for Google and “DS” for D-Search in all the following charts. The group-oriented results are analysed by using one-way ANOVA and the system-oriented results are analysed by using t-test.

6.2.1.1 Group-oriented Results

Table 6.7 shows the statistical results of task completion time for the novice user group in performing simple tasks where the first table gives descriptives and the second table lists one-way ANOVA results. The best mean value was found on D-Search (20.79 seconds) followed by the Open Directory (115.65 seconds) and Google (178.75 seconds). The minimum user task completion time was found on D-Search (5.53 seconds) while the maximum time was found on Google (464.94 seconds). The minimum standard deviation was found on D-Search (17.67 seconds) followed by the Open Directory (59.39 seconds) and Google (178.19 seconds). There is significant difference among the time of three systems for novice users (P = 0.028).

System	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
OD	8	115.65	59.39	21.00	66.00	165.30	27.64	224.90
DS	8	20.79	17.67	6.25	6.02	35.57	5.53	56.60
G	8	178.74	178.19	63.00	29.77	327.70	42.50	464.94
Total	24	105.06	123.41	25.19	52.95	157.17	5.53	464.94

System	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	101131.01	2	50565.50	**4.262**	**.028**
Within Groups	249134.07	21	11863.53		
Total	350265.07	23			

Table 6.7 The Novice Group's Task Completion Time Results for Simple Tasks (P<0.05)

Table 6.8 shows the statistical results of task completion time for the novice user group in performing complex tasks where the first table gives descriptives and the second table lists one-way ANOVA results. The smallest mean value was found on Google (143.51 seconds) followed by D-Search (143.51 seconds) and the Open Directory (230.98 seconds). The minimum user task completion time was found on Google (63.40 seconds) while the maximum time was found on the Open Directory (513.77 seconds).

The minimum SD was found on Google (120.80 seconds) followed by D-Search (63.50 seconds) and the Open Directory (170.02 seconds). There is no significant difference among the time of three systems for novice users (P = 0.119).

System	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
OD	8	230.98	170.02	60.11	88.84	373.12	67.70	513.77
DS	8	143.51	63.50	22.45	90.42	196.60	79.07	233.45
G	8	120.80	38.31	13.55	88.78	152.83	63.40	167.93
Total	24	165.10	113.25	23.12	117.27	212.92	63.40	513.77

System	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	54147.23	2	27073.614	**2.361**	**.119**
Within Groups	240851.33	21	11469.11		
Total	294998.56	23			

Table 6.8 The Novice Group's Task Completion Time Results for Complex Tasks (P>0.05)

Figure 6.1 illustrates a summary of mean and SD from Table 6.7 and Table 6.8 in terms of task complexity. Unlike D-Search and the Open Directory where both mean and SD of the task completion time on these systems rose noticeably with the growth of task complexity, the task completion time on Google dropped a little. The better performance shown on Google reflected the group's weakness in using search strategies. This is because unlike Google, the Open Directory could only provide a direction instead of an exact location in obtaining extra information on the websites. That is, it tells a user the websites listed in this category (e.g., “Europe: United Kingdom: Business and Economy: Shopping: Photography and Optics”) are very likely to sell these kinds of products the user is looking for but it will not tell which websites in this list are selling the specific products they are looking for. In comparison, if the user types a specific product name in Google, Google will return websites selling this product. Thus, the user needs more time and effort to get results in the directory if they are not experienced in deciding a good starting point before using the search service in the directory.

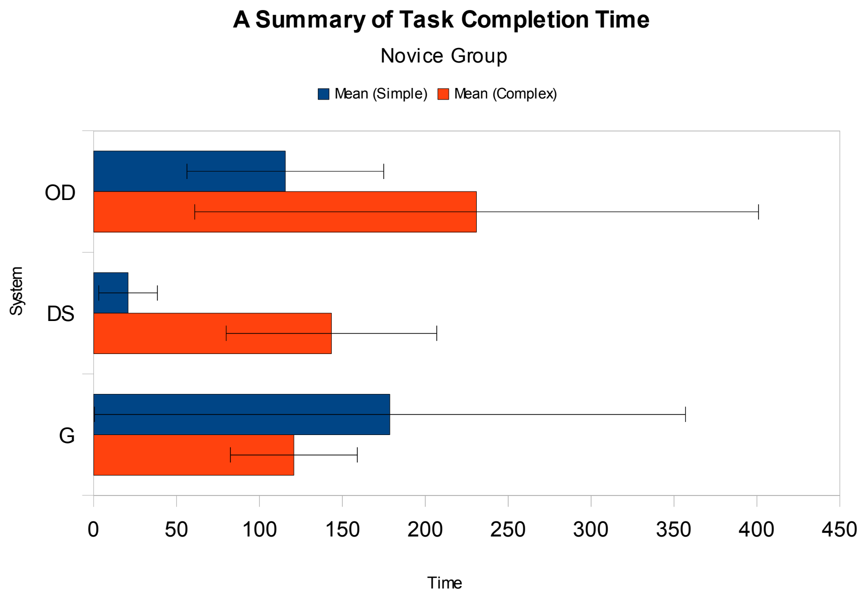


Figure 6.1 A Summary of the Novice Group's Task Completion Time

Table 6.9 shows the statistical results of task completion time for the expert user group in performing simple tasks where the first table gives descriptives and the second table lists one-way ANOVA results. The minimum mean was found on D-Search (11.48 seconds) followed by the Open Directory (55.89 seconds) and Google (119.66 seconds). The minimum user task completion time was found on D-Search (3.50 seconds) while the maximum time was found on Google (396.85 seconds). The minimum value of SD was found on D-Search (6.28 seconds) while the maximum value of SD was found on Google (131.69 seconds). There is no significant difference among the time of three systems for expert users (P = 0.057).

System	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
OD	8	55.89	64.29	22.73	2.15	109.63	15.85	208.80
DS	8	11.48	6.28	2.22	6.23	16.73	3.50	20.70
G	8	119.66	131.69	46.56	9.55	229.76	24.50	396.85
Total	24	62.35	92.76	18.94	23.17	101.52	3.50	396.85

System	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	47313.95	2	23656.97	**3.299**	**.057**
Within Groups	150600.24	21	7171.44		
Total	197914.19	23			

Table 6.9 The Expert Group's Task Completion Time Results for Simple Tasks (P>0.05)

Table 6.10 shows the statistical results of task completion time for the expert user group in performing complex tasks where the first table gives descriptives and the second table lists one-way ANOVA results. The best mean value was found on the Open Directory (78.40 seconds) followed by D-Search (91.98 seconds) and Google (97.33 seconds). The minimum user task completion time was found on the Open Directory (15.37 seconds) while the maximum time was found on D-Search (172.33 seconds). The minimum value of SD was found on Google (26.19 seconds) while the maximum value of SD was found on D-Search (43.64 seconds).There is no significant difference found among the time of three systems in the expert user group (P = 0.602).

System	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
OD	8	78.40	42.52	15.03	42.86	113.95	15.37	132.01
DS	8	91.98	43.64	15.43	55.50	128.47	37.32	172.33
G	8	97.33	26.19	9.26	75.43	119.22	69.19	156.56
Total	24	89.24	37.48	7.65	73.41	105.07	15.37	172.33

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1522.94	2	761.47	**.519**	**.602**
Within Groups	30789.63	21	1466.17		
Total	32312.57	23			

Table 6.10 The Expert Group's Task Completion Time Results for Complex Tasks (P>0.05)

Figure 6.2 illustrates a summary of mean and SD from Table 6.9 and Table 6.10 in terms of task complexity which shows a similar trend as in the novice group.

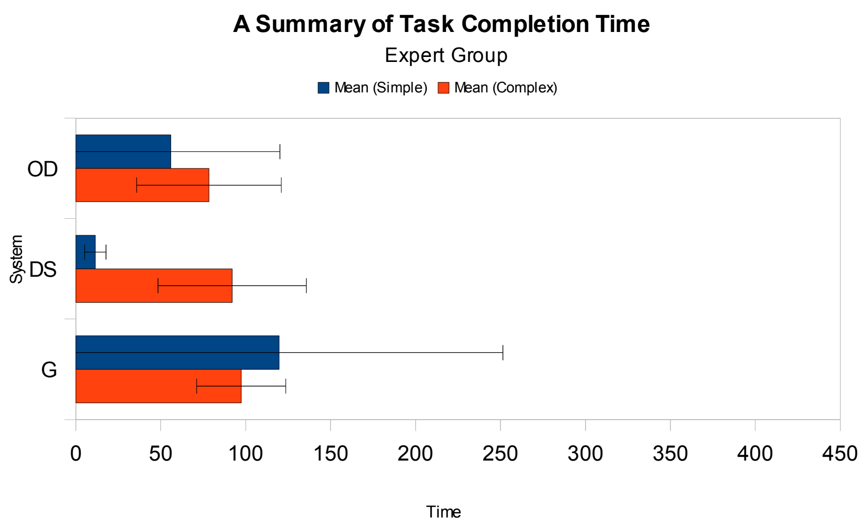


Figure 6.2 A Summary of the Expert Group's Task Completion Time

6.2.1.2 System-oriented Results

Table 6.11 shows the statistical results of task completion time of two user groups for completing simple tasks on the Open Directory where *t*-test was used. The expert user group reported better mean value (55.89 seconds) than the novice group (115.65 seconds) while the novice group reported slightly better SD value (59.39 seconds) than the expert group (64.29 seconds). There is no significant difference found between two groups (P = 0.074).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	115.6500	59.3900	21.0000	t = -1.891
Expert	8	55.8900	64.2900	22.7300	P = 0.662

Table 6.11 The Open Directory's Task Completion Time Results for Simple Tasks (P>0.05)

Table 6.12 shows the statistical results of task completion time of two user groups for completing complex tasks on the Open Directory where *t* and *P* were calculated by using *t*-test. The expert user group reported better mean (78.40 seconds) and SD (42.52 seconds) than the novice group (230.98 seconds and 170.02 seconds). There is significant difference found between two groups (P = 0.040).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	230.9800	170.0200	60.1100	t = -2.462
Expert	8	78.4000	42.5200	15.0300	P = 0.040

Table 6.12 The Open Directory's Task Completion Time Results for Complex Tasks (P<0.05)

Figure 6.3 illustrates a summary of mean and SD from Table 6.11 and Table 6.12 where significant growth was shown in both groups when the task complexity rose.

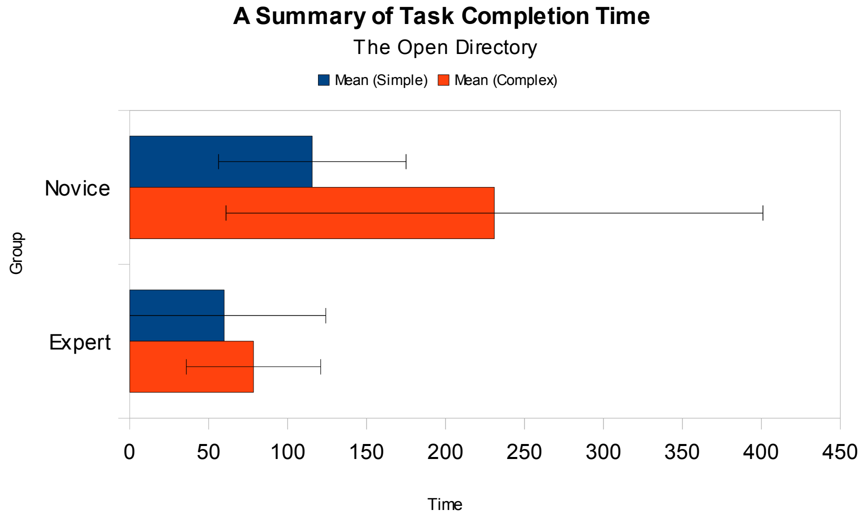


Figure 6.3 A Summary of the Open Directory's Task Completion Time

Table 6.13 shows the statistical results of task completion time of two user groups for completing simple tasks on D-Search where *t*-test was used. The expert user group reported better mean and SD than the novice group. There is no significant difference found between two groups (P = 0.195).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	20.7900	17.6700	6.2500	t = -1.404
Expert	8	11.4800	6.2800	2.2200	P = 0.195

Table 6.13 D-Search's Task Completion Time Results for Simple Tasks (P>0.05)

Table 6.14 shows the statistical results of task completion time of two user groups for completing complex tasks on D-Search where *t*-test was used. A similar trend as for completing simple tasks was found. There is also no significant difference observed between two groups (P = 0.662).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	143.5100	63.5000	22.4500	t = -1.891
Expert	8	91.9800	43.6400	15.4300	P = 0.662

Table 6.14 D-Search's Task Completion Time Results for Complex Tasks (P>0.05)

Figure 6.4 shows an overview of mean and SD from Table 6.13 and Table 6.14 where a similar trend of time growth was found as on the Open Directory when the task complexity rose.

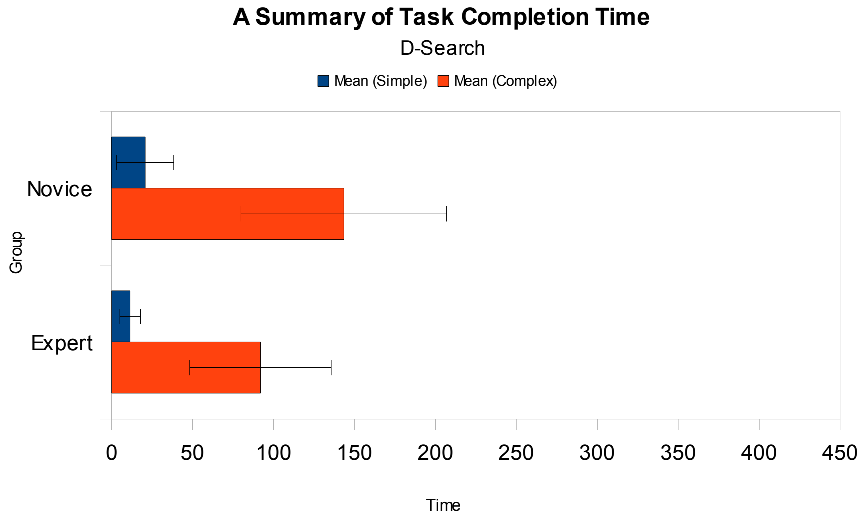


Figure 6.4 A Summary of Task Completion Time on D-Search

Table 6.15 shows the statistical results of task completion time of two user groups for completing simple tasks on Google where *t*-test was used. The expert user group reported slightly better mean and SD than the novice group. There is no significant difference found between two groups (P = 0.463).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	178.7400	178.1900	63.0000	$t = -0.754$
Expert	8	119.6600	131.6900	46.5600	$P = 0.463$

Table 6.15 Google's Task Completion Time Results for Simple Tasks (P>0.05)

Table 6.16 shows the statistical results of task completion time of two user groups for completing complex tasks on Google where *t*-test was used. A similar trend as for completing simple tasks was found. There is also no significant difference observed between two groups (P = 0.662).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	120.8000	38.3100	13.5500	$t = -1.431$
Expert	8	97.3300	26.1900	9.2600	$P = 0.174$

Table 6.15 Google's Task Completion Time Results for Complex Tasks (P>0.05)

Figure 6.5 shows an overview of mean and SD from Table 6.15 and Table 6.16 where both groups reported slightly completion time drop when the task complexity rose.

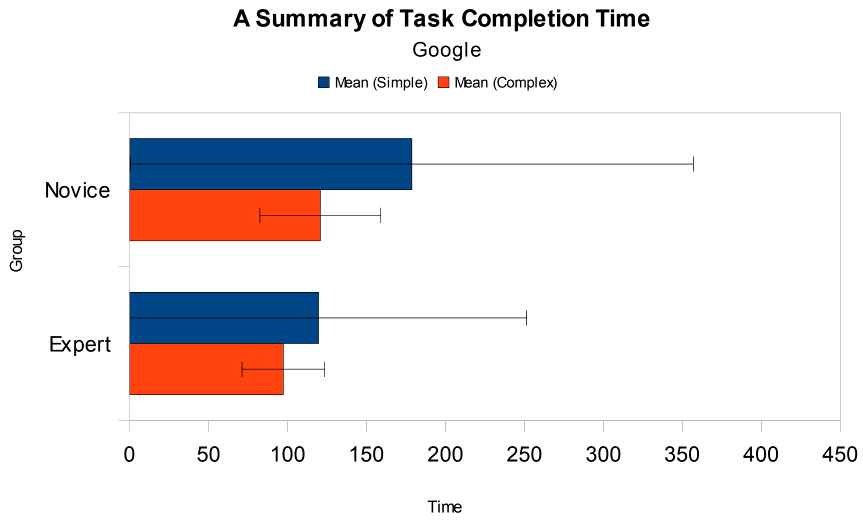


Figure 6.5 A Summary of Google's Task Completion Time

A summary of the results in this section is given in Table 6.17.

	Simple		T-test (p value)	Complex		T-test (p value)
	Novice	Expert		Novice	Expert	
OD	115.65	55.89	0.662	230.98	78.4	0.040
DS	20.79	11.48	0.195	143.51	91.98	0.662
G	178.74	119.66	0.463	120.8	97.33	0.174
One-way ANOVA (p value)	0.028	0.057		0.119	0.602	

Table 6.17 A Summary of Task Completion Time

6.2.2 Normalised Success Rate

Normalised success rate uses the users' session success rate to normalise their overall task success rate in order to acquire more accurate understandings for the effectiveness of comparative systems. Results for the normalised success rate are analysed by using Fisher's Exact Test but they are explained in the same way as the previous section.

6.2.2.1 Group-oriented Results

Table 6.18 shows the statistical results of normalised success rate for the novice user group in performing simple tasks on the three systems where X^2 statistical test for contingency table was used for calculating the P-value. The lowest success rate was found on the Open Directory (0.090) while the best success rate are found on D-Search (0.854) and Google (0.875). There is significant difference found among the systems (P = 0.008).

	Task Output		Total	Rate	
	Success	Fail			
OD	4	4	8	0.090	*P* = 0.008
DS	8	0	8	0.854	
G	8	0	8	0.875	
Total	20	4	24		

Table 6.18 The Novice Group's Normalised Success Rate for Simple Tasks (P<0.05)

Table 6.19 shows the statistical results of normalised success rate for the novice user group in performing simple tasks on the three systems where X^2 statistical test for contingency table was used for calculating the P-value. The lowest success rate was found on the Open Directory (0.067) while the best success rate was found on D-Search (0.940) and Google (0.670). Note Google's rate drop was caused by a failed task. There is also significant difference found among these systems (P = 0.000).

	Task Output		Total	Rate	
	Success	Fail			
OD	1	7	8	0.067	*P* = 0.000
DS	8	0	8	0.875	
G	7	1	8	0.670	
Total	16	8	24		

Table 6.19 The Novice Group's Normalised Success Rate for Complex Tasks (P<0.05)

Figure 6.6 illustrates a summary of normalised success rate for the novice user group. The rate remained nearly the same on D-Search with the growth of the task complexity whereas it dropped slightly on the Open Directory and Google.

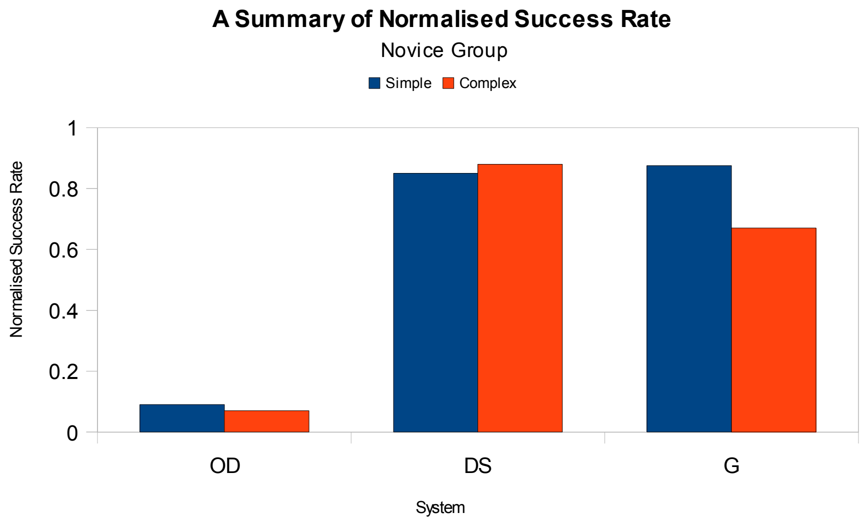


Figure 6.6 A Summary of the Novice Group's Normalised Success Rate

Table 6.20 shows the statistical results of normalised success rate for the expert user group in performing simple tasks on the three systems where X^2 statistical test for contingency table was used for calculating the P-value. A similar trend of rate performance as for the novice group was found for the expert group. There is significant difference among these systems (P = 0.032).

	Task Output		Total	Rate	
	Success	Fail			
OD	5	3	8	0.260	$P = 0.032$
DS	8	0	8	0.875	
G	8	0	8	0.875	
Total	21	3	24		

Table 6.20 The Expert User Group's Normalised Success Rate for Simple Tasks (P<0.05)

Table 6.21 shows the statistical results of normalised success rate for the expert user group in performing simple tasks on the three systems where X^2 statistical test for contingency table was used for calculating the P-value. A similar trend of rate performance as for the simple tasks was found. There is significant difference among these systems (P = 0.002).

	Task Output		Total	Rate	
	Success	Fail			
OD	2	6	8	0.113	$P = 0.002$
DS	8	0	8	0.917	

G	7	1	8	0.875	
Total	17	7	24		

Table 6.21 The Expert User Group's Normalised Success Rate for Complex Tasks (P<0.05)

Figure 6.7 illustrates a summary of normalised success rate for the expert group. The rate was slightly dropped on the Open Directory when it remained more or less the same on D-Search and Google.

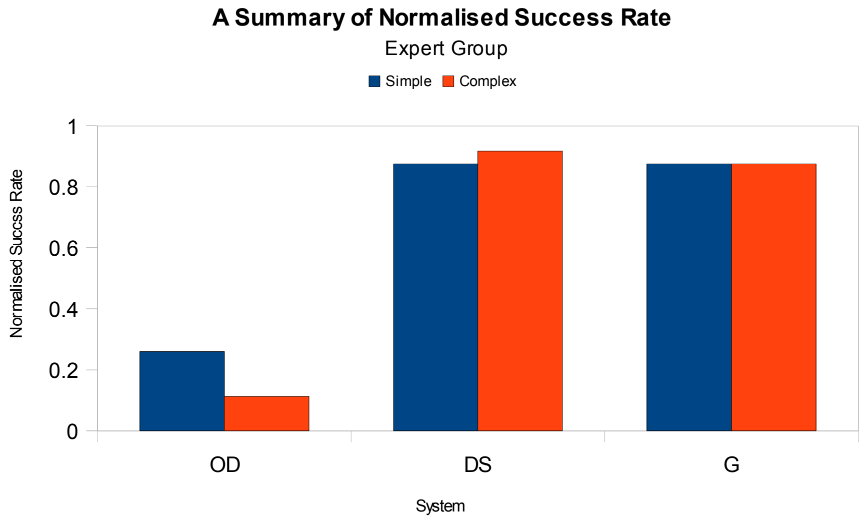


Figure 6.7 A Summary of the Expert User Group's Normalised Success Rate

6.2.2.2 System-oriented Results

Table 6.22 shows the statistical results of normalised success rate of two user groups for completing simple tasks on the Open Directory where *Fisher's Exact Test* was used. Both groups performed poorly on the Open Directory although the expert user group reported slightly better normalised success rate (0.260) than the novice group (0.090). There is no significant difference found between two groups (P = 0.343).

	Task Output		Rate	
	Success	**Fail**		
Novice	4	4	0.090	*P* = 0.343
Expert	5	3	0.260	
Total	9	7		

Table 6.22 The Open Directory's Normalised Success Rate for Simple Tasks (P>0.05)

Table 6.23 shows the statistical results of normalised success rate of two user groups for completing complex tasks on the Open Directory where *Fisher's Exact Test* was used. With the growth of task complexity, both groups performed even more poorly on the Open Directory but similarly to performing simple tasks, the expert group reported better normalised success rate (0.113) than the novice group (0.067). There is no significant difference found between two groups (P = 0.400).

	Task Output		**Rate**	
	Success	**Fail**		
Novice	1	7	0.113	*P* = 0.400
Expert	2	6	0.067	
Total	3	13		

Table 6.23 The Open Directory's Normalised Success Rate for Complex Tasks (P>0.05)

Figure 6.8 illustrates a summary of normalised success rate from Table 6.22 and Table 6.23 where poor rate was shown in each group and significant drop was also shown in each group when the task complexity rose.

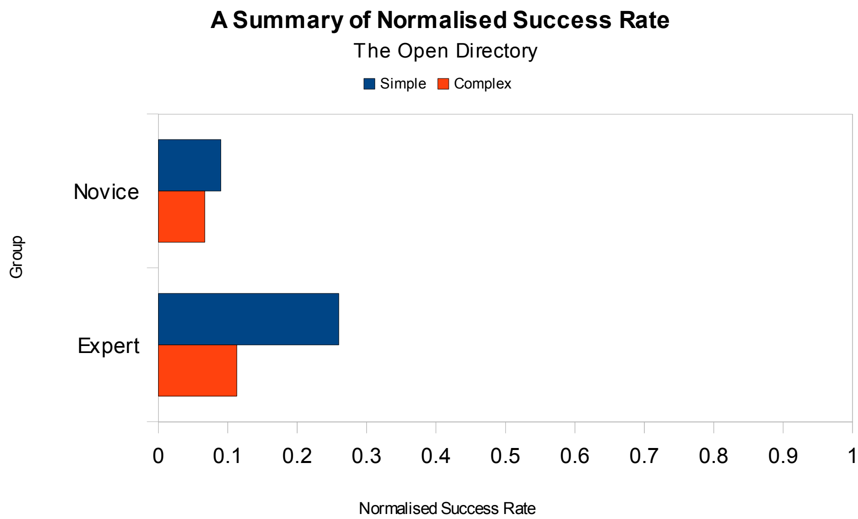


Figure 6.8 A Summary of the Open Directory's Normalised Success Rate

Table 6.24 shows the results of normalised success rate of two user groups for completing simple tasks on D-Search where no statistical test was used due to the 100% task success rate. Both groups reported good and close success rate with D-Search. There is no significant difference found between the two groups (P = 1.000).

	Task Output	**Rate**	*P* = 1.000

	Success	Fail		
Novice	8	0	0.854	
Expert	8	0	0.875	
Total	16	0		

Table 6.24 D-Search's Normalised Success Rate for Simple Tasks (P > 0.05)

Table 6.25 shows the results of normalised success rate of two user groups for completing complex tasks on D-Search where no statistical test was used due to the 100% task success rate. Both groups reported good and close success rate with D-Search. There is no significant difference found between the two groups (P = 1.000).

	Task Output		**Rate**	
	Success	**Fail**		
Novice	8	0	0.875	*P* = 1.000
Expert	8	0	0.917	
Total	16	0		

Table 6.25 D-Search's Normalised Success Rate for Complex Tasks (P > 0.05)

Figure 6.9 illustrates a summary of normalised success rate from Table 6.24 and Table 6.25 where good rate was shown in each group and an interesting rate rise was also shown in each group when the task complexity grew.

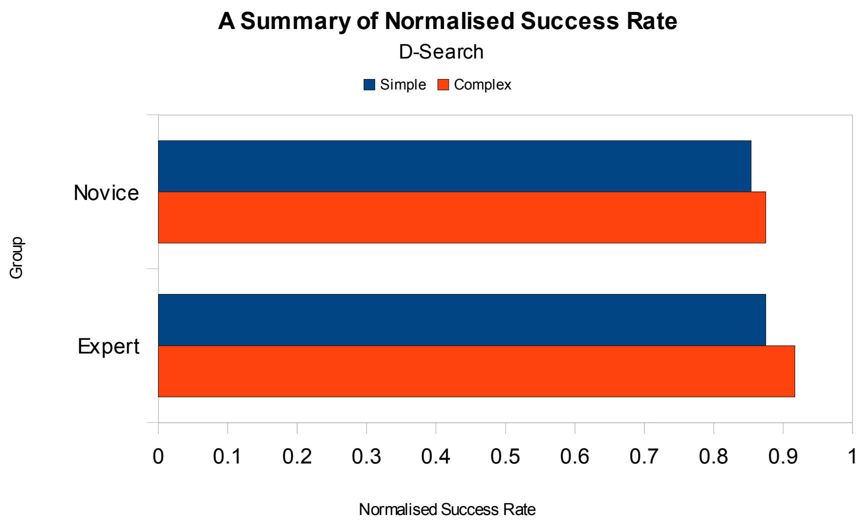


Figure 6.9 A Summary of D-Search's Normalised Success Rate

Table 6.26 shows the results of normalised success rate of two user groups for completing simple tasks on Google where no statistical test was used due to the 100% task success rate. Both groups reported same good success rate with Google. There is no significant difference found between the two groups (P = 1.000).

	Task Output		Rate	
	Success	**Fail**		
Novice	8	0	0.875	*P* = 1.000
Expert	8	0	0.875	
Total	16	0		

Table 6.26 Google's Normalised Success Rate for Simple Tasks (P > 0.05)

Table 6.27 shows the results of normalised success rate of two user groups for completing complex tasks on Google where *Fisher's Exact Test* was used. Both groups reported good success rate with Google and the expert group reported better rate (0.875) than the novice group (0.670) although the same number of tasks was failed. There is no significant difference found between the two groups (P = 0.533).

	Task Output		Rate	
	Success	**Fail**		
Novice	7	1	0.670	*P* = 0.533
Expert	7	1	0.875	
Total	14	2		

Table 6.27 Google's Normalised Success Rate for Complex Tasks (P > 0.05)

Figure 6.10 illustrates a summary of normalised success rate from Table 6.26 and Table 6.27 where good rate was shown in each group and the novice group's rate dropped due to the growth of task complexity.

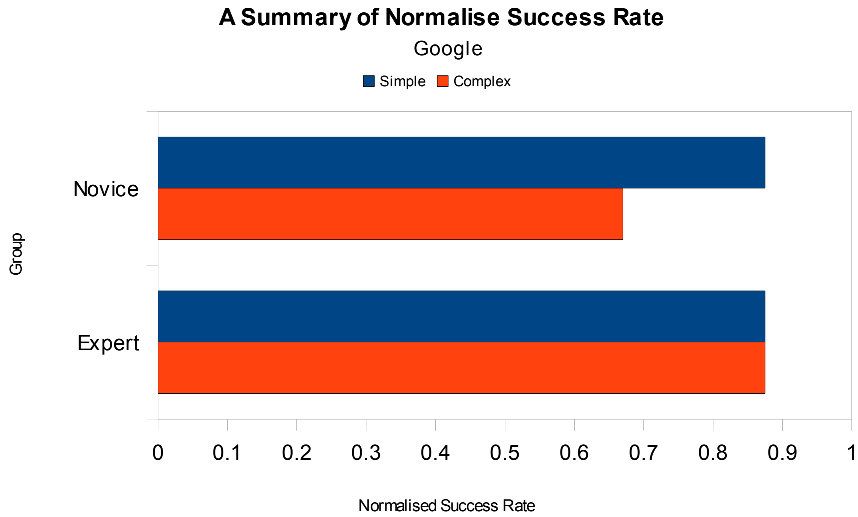


Figure 6.10 A Summary of Google's Normalised Success Rate

A summary of the results in this section is given in Table 6.28.

	Simple		Fisher's Exact	Complex		Fisher's Exact
	Novice	Expert	Test (p value)	Novice	Expert	Test (p value)
OD	0.090	0.260	0.343	0.113	0.067	0.400
DS	0.854	0.875	1.000	0.875	0.917	1.000
G	0.875	0.875	1.000	0.670	0.875	0.533
X^2 (p-value)	0.008	0.000		0.032	0.002	

Table 6.28 A Summary of Normalised Success Rate

6.2.3 User Pathway Data

User pathway data provide details for the system performance based on the average number of queries participants used in designated tasks on the system. Similar to Task Completion Time, group-oriented results are analysed by using one-way ANOVA and system-oriented results are analysed by using t-test.

6.2.3.1 Group-oriented Results

Table 6.29 shows the statistical results of query usage for the novice user group in performing simple tasks where the first table gives descriptives and the second table lists one-way ANOVA results. The mean values of D-Search and Google (1.3750 and 1.6250 respectively) significantly outperformed the Open Directory (7.0000), which indicated the average queries used by the novice user group on each system. The minimum queries used on each system were similar whereas the maximum queries used on the Open Directory (19.00) was much more than D-Search (3.00) and Google (3.00). There is significant difference among the query usage of three systems for novice users ($P = 0.007$).

System	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
OD	8	7.0000	6.1179	2.1630	1.8853	12.1147	2.00	19.00
DS	8	1.3750	0.7440	0.2631	0.7530	1.9970	1.00	3.00
G	8	1.6250	0.7440	0.2631	1.0030	2.2470	1.00	3.00
Total	24	3.3333	4.3306	0.8840	1.5047	5.1620	1.00	19.00

System	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	161.583	2	80.790	**6.290**	**.007**
Within Groups	269.750	21	12.850		
Total	431.333	23			

Table 6.29 The Novice Group's Query Usage Results for Simple Tasks (P<0.05)

Table 6.30 shows the statistical results of query usage for the novice user group in performing complex tasks where the first table gives descriptives and the second table lists one-way ANOVA results. Similar to simple tasks, the mean value which indicated the average queries used by the novice user group on each system reported that D-Search and Google (1.1250 and 2.3750 respectively) significantly outperformed the Open Directory (11.6250). The minimum queries used on each system were the same (1.00) whereas the maximum queries used on the Open Directory (25.00) was much more than D-Search (2.00) and Google (5.00). There is significant difference among the query usage of three systems for novice users (P = 0.000).

System	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
OD	8	11.6250	7.8729	2.7835	5.0431	18.2069	1.00	25.00
DS	8	1.1250	0.3536	0.1250	0.8294	1.4206	1.00	2.00
G	8	2.3750	1.5059	0.5324	1.1160	3.6340	1.00	5.00
Total	24	5.0417	6.5174	1.3304	2.2896	7.7937	1.00	25.00

System	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	526.333	2	263.167	**12.264**	**.000**
Within Groups	450.625	21	21.458		
Total	976.958	23			

Table 6.30 The Novice Group's Query Usage Results for Complex Tasks (P<0.05)

Figure 6.11 illustrates a summary of mean and SD from Table 6.29 and Table 6.30 in terms of query usage. Open Directory showed dramatical query rise with the growth of task complexity while changes on D-Search and Google looked more stable although D-Search tended to outperform Google a little.

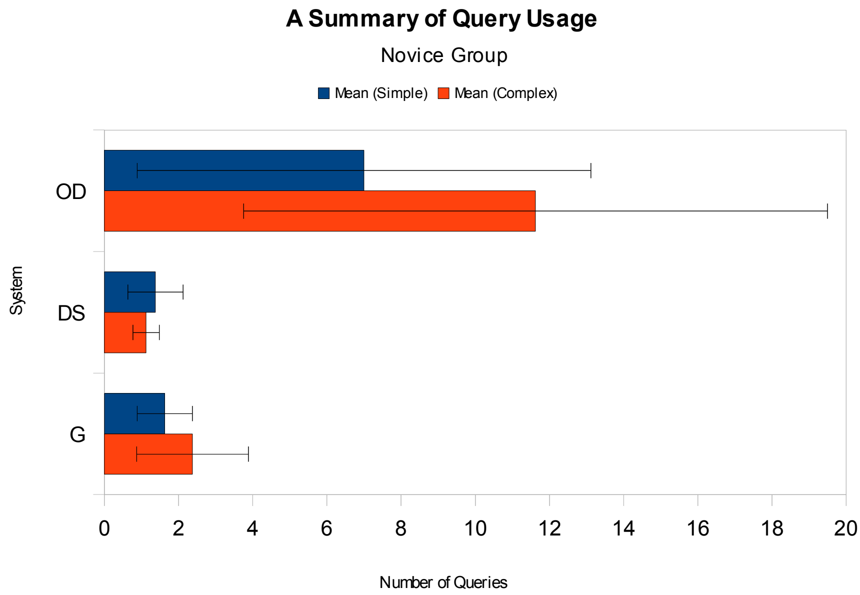


Figure 6.11 A Summary of the Novice Group's Query Usage

Table 6.31 shows the statistical results of query usage for the expert user group in performing simple tasks where the first table gives descriptives and the second table lists one-way ANOVA results. The average number of queries used on D-Search and Google (1.2500 and 1.7500 respectively) were slightly better than the Open Directory (3.1250). The minimum queries used on each system were the same whereas the maximum queries used on the Open Directory (5.00) was higher than D-Search (2.00) and Google (3.00). There is significant difference among the query usage of three systems for novice users (P = 0.002).

System	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
OD	8	3.1250	1.3562	0.4795	1.9912	4.2588	1.00	5.00
DS	8	1.2500	0.4629	0.1637	0.8630	1.6370	1.00	2.00
G	8	1.7500	0.7071	0.2500	1.1588	2.3412	1.00	3.00
Total	24	2.0417	1.1971	0.2444	1.5362	2.5471	1.00	5.00

System	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	15.083	2	7.542	**8.860**	**.002**
Within Groups	17.875	21	0.851		
Total	32.958	23			

Table 6.31 The Expert Group's Query Usage Results for Simple Tasks (P<0.05)

Table 6.32 shows the statistical results of query usage for the expert user group in performing complex tasks where the first table gives descriptives and the second table lists one-way ANOVA results. Similar to simple tasks, the mean value which indicated the average queries used by the expert user group on each system reported that D-Search and Google (1.2500 and 1.5000 respectively) significantly outperformed the Open Directory (8.6250). The minimum queries used on each system were similar whereas the maximum queries used on the Open Directory (18.00) was much more than D-Search (3.00) and Google (2.00). There is significant difference among the query usage of three systems for novice users (P = 0.000).

System	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
OD	8	8.6250	5.5016	1.9451	4.0255	13.2245	2.00	18.00
DS	8	1.2500	0.7071	0.2500	0.6588	1.8412	1.00	3.00
G	8	1.5000	0.5345	0.1890	1.0531	1.9469	1.00	2.00
Total	24	3.7917	4.6530	0.9498	1.8269	5.7565	1.00	18.00

System	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	280.583	2	140.292	**13.553**	**.000**
Within Groups	217.375	21	10.351		
Total	497.958	23			

Table 6.32 The Expert Group's Query Usage Results for Complex Tasks (P<0.05)

Figure 6.12 illustrates a summary of mean and SD from Table 6.31 and Table 6.32 in terms of query usage. Similar to novice group, Open Directory showed dramatical query rise with the growth of task complexity while changes on D-Search and Google were insignificant although D-Search slightly outperformed Google.

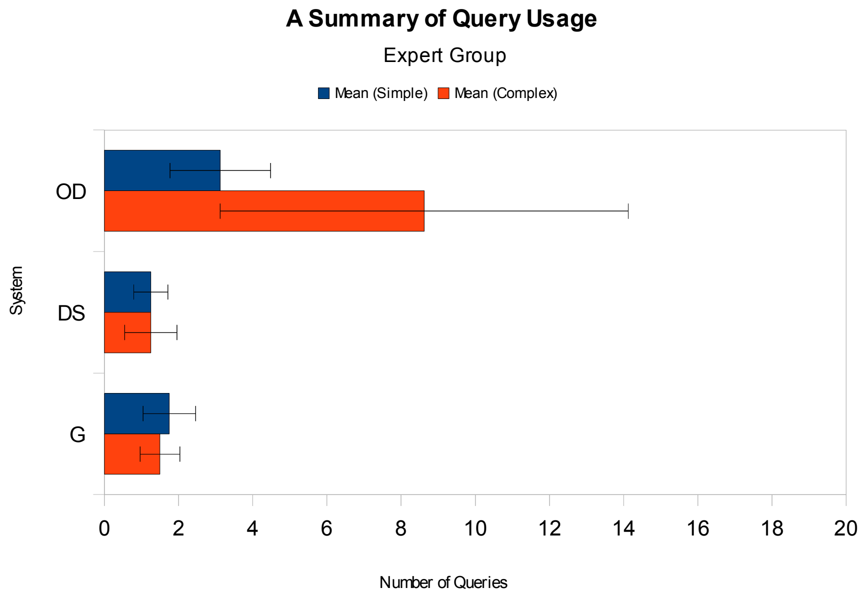


Figure 6.12 A Summary of the Expert Group's Query Usage

6.2.3.2 System-oriented Results

Table 6.33 shows the statistical results of query usage of two user groups for completing simple tasks on the Open Directory where *t*-test was used. The expert user group reported better mean (3.1250) and SD (1.3562) than the novice group (7.0000 and 6.1179). There is no significant difference found between two groups (P = 0.120).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	7.0000	6.1179	2.1630	$t = -1.749$
Expert	8	3.1250	1.3562	0.4795	$P = 0.120$

Table 6.33 The Open Directory's Query Usage for Simple Tasks (P>0.05)

Table 6.34 shows the statistical results of query usage of two user groups for completing complex tasks on the Open Directory where *t* and *P* were calculated by using *t*-test. Again, the expert user group reported better results that the novice group in both Mean and SD. There is no significant difference found between two groups (P = 0.392).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	11.6250	7.8729	2.7835	$t = -0.883$
Expert	8	8.6250	5.5016	1.9451	$P = 0.392$

Table 6.34 The Open Directory's Query Usage for Complex Tasks (P>0.05)

Figure 6.13 illustrates a summary of mean and SD from Table 6.33 and Table 6.34 where significant growth was shown in both groups when the task complexity rose. In addition, the expert group performed generally better than the novice group regardless task complexity due to its better experience.

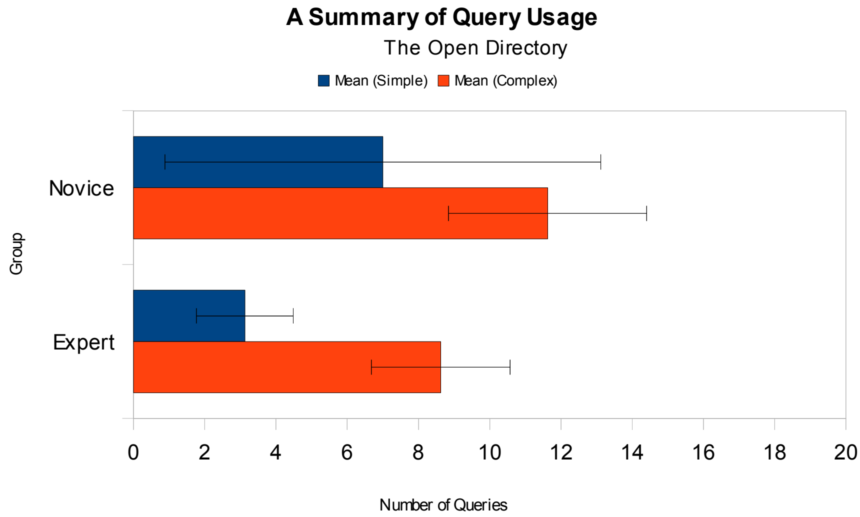


Figure 6.13 A Summary of the Open Directory's Query Usage

Table 6.35 shows the statistical results of query usage of two user groups for completing simple tasks on D-Search where *t*-test was used. The two groups reported similar mean and SD while the expert user group showed slightly better results in both. There is no significant difference found between two groups (P = 0.693).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	1.2500	0.7440	0.2631	t = -0.403
Expert	8	1.3750	0.4629	0.1637	P = 0.693

Table 6.35 D-Search's Query Usage for Simple Tasks (P>0.05)

Table 6.36 shows the statistical results of query usage of two user groups for completing complex tasks on D-Search where t and P were calculated by using *t*-test. Again, the two groups reported similar mean and SD. There is no significant difference found between two groups (P = 0.662).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	1.1250	0.3536	0.1250	t = -0.447
Expert	8	1.2500	0.7071	0.2500	P = 0.662

Table 6.36 D-Search's Query Usage for Complex Tasks (P>0.05)

Figure 6.14 illustrates a summary of mean and SD from Table 6.35 and Table 6.36 where none significant difference was found in both groups which indicated the stable performance of D-Search regardless the complexity of tasks and group difference.

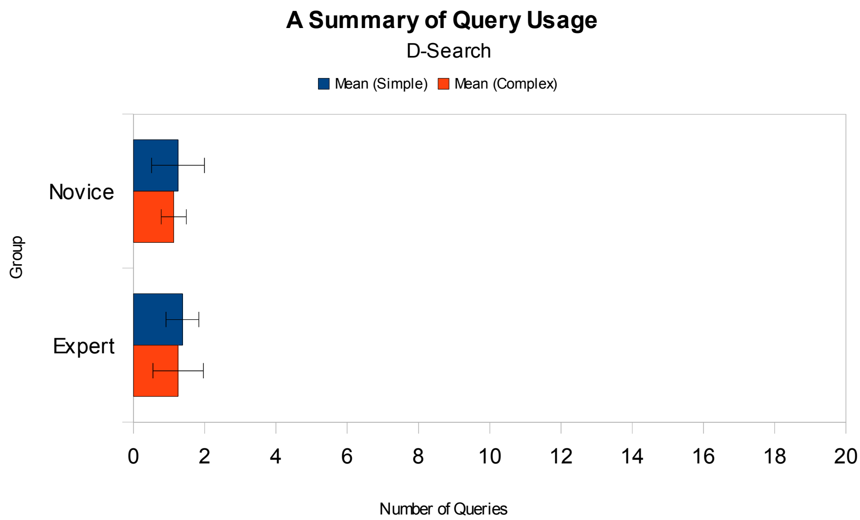


Figure 6.14 A Summary of D-Search's Query Usage

Table 6.37 shows the statistical results of query usage of two user groups for completing simple tasks on Google where *t*-test was used. The two groups reported similar mean and SD while the expert user group showed slightly better results in both. There is no significant difference found between two groups (P = 0.736).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	1.6250	0.7440	0.2631	*t* = -0.344
Expert	8	1.7500	0.7071	0.2500	*P* = 0.736

Table 6.37 Google's Query Usage for Simple Tasks (P>0.05)

Table 6.38 shows the statistical results of query usage of two user groups for completing complex tasks on D-Search where *t* and *P* were calculated by using *t*-test. Again, the two groups reported similar mean and SD. There is no significant difference found between two groups (P = 0.144).

Group	N	Mean	Std. Deviation	Std. Error Mean	
Novice	8	2.3750	1.5059	0.5324	*t* = -1.549
Expert	8	1.5000	0.5345	0.1890	*P* = 0.144

Table 6.38 Google's Query Usage for Complex Tasks (P>0.05)

Figure 6.15 illustrates a summary of mean and SD from Table 6.37 and Table 6.38

where the difference between groups for tasks having the same complexity and the difference in each group when the task complexity rose were insignificant. This also indicated the stable performance of Google regardless the complexity of tasks and group difference.

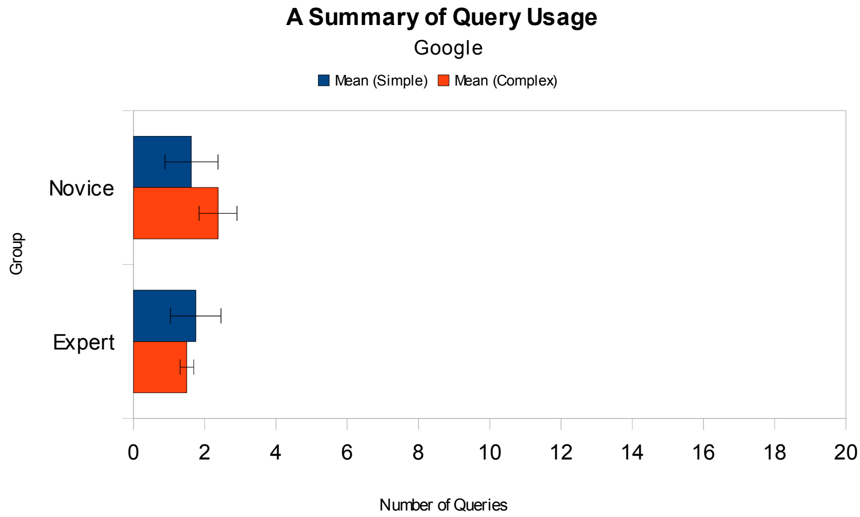


Figure 6.15 A Summary of Google's Query Usage

A summary of the results in this section is given in Table 6.39.

	Simple		T-test (p value)	Complex		T-test (p value)
	Novice	Expert		Novice	Expert	
OD	7.000	3.125	0.120	11.625	8.625	0.392
DS	1.250	1.375	0.69	1.125	1.250	0.66
G	1.625	1.750	0.74	2.375	1.500	0.144
One-way ANOVA (p value)	0.007	0.000	-	0.002	0.000	-

Table 6.39 A Summary of User Pathway Data

6.2.4 User Feedback Ratings

A 1 - 6 rating system[18] was used to capture user feedback in a questionnaire covering assessments on the effectiveness (i.e., success), efficiency (i.e., quickness), satisfaction (i.e., satisfaction of user experience) and search helpfulness (i.e., help in user

18 In the 1 – 6 satisfaction rating system, from the maximum rating to the minimum, 6 represents the highest positive response and 1 indicates the highest negative response to a statement. In details, 1 – strongly disagree, 2 – disagree, 3 – slightly disagree, 4 – slightly agree, 5 – agree and 6 – strongly agree.

navigation). Each set of assessments include three statements. Users were firstly asked to give standalone ratings on a single statement about D-Search. For example, “with D-Search, you can find your desired results successfully”. Then they were asked to give ratings for the same topic in comparison to Google or the Open Directory. For example, “you can find your desired results with D-Search more successfully than with Google”. Results in this section are only produced in group-oriented views with one-way ANOVA test.

6.2.4.1 Effectiveness Ratings

For the question that “I found results successfully with D-Search (1 strongly disagree – 6 strongly agree)”, the detailed user ratings given by the two groups are shown in the first table in Table 6.40 and the relevant statistical results are shown in the second table where t-test was used. The two groups reported same mean value (5.0000) which indicated a very positive response to the effectiveness of D-Search. There is no significant difference between the two groups (P = 1.000) which indicated that the positive ratings were system-independent rather than group-independent.

Group	N	Distribution						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	0	0	4	4	4	0.00	0.00	0.00	33.33	33.33	33.33
Expert	12	0	0	1	1	7	3	0.00	0.00	8.33	8.33	58.33	33.33

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	5.0000	0.8528	0.2462	0.000	1.000
Expert	5.0000	0.8528	0.2462		

Table 6.40 Independent User Effectiveness Ratings of D-Search (P > 0.05)

For the question that “I found results more successfully with D-Search than the Open Directory (1 strongly disagree – 6 strongly agree)”, the detailed ratings given by the two groups are shown in Table 6.41. The two groups reported similar mean value (> 5.0000) which indicated very positive response to the effectiveness of D-Search against the Open Directory. Although the expert group gave slightly higher average ratings, there is no significant difference between the two groups (P = 0.839).

Group	N	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	0	0	2	4	6	0.00	0.00	0.00	16.67	33.33	50.00
Expert	12	0	1	0	0	3	8	0.00	8.33	0.00	0.00	25.00	66.67

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	5.3333	0.7785	0.2247	0.206	0.839
Expert	5.4167	1.1645	0.3362		

Table 6.41 User Effectiveness Ratings of D-Search Compared to The Open Directory (P > 0.05)

For the question that "I found results more successfully with D-Search than Google (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in Table 6.42. The two groups reported similar mean value between 4.0000 and 5.0000 which indicated slightly positive response to the effectiveness of D-Search against Google. Similarly, although the expert group gave slightly higher average ratings, there is no significant difference between the two groups (P = 0.307).

Group	Size	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	0	3	3	5	1	0.00	0.00	25.00	25.00	41.67	8.33
Expert	12	0	0	1	4	4	3	0.00	0.00	8.33	33.33	33.33	25.00

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	4.3333	0.9847	0.2843	1.047	0.307
Expert	4.7500	0.9653	0.2787		

Table 6.42 User Effectiveness Ratings of D-Search Compared to Google (P > 0.05)

Figure 6.16 presents an overview of the mean and SD from Table 6.40 – 6.42 where showed overall positive ratings and rating trends by both groups. The expert group always gave more positive ratings than the novice group. Both groups tended to give higher effectiveness ratings when comparing D-Search and the Open Directory whereas they tended to lower the ratings when comparing to Google. Note in the figure, STD stands for independent D-Search relevant question, VOD stands for comparative D-Search question against the Open Directory and VG stands for comparative D-Search question against Google. This notation is also used in the following sub-sections from 6.2.4.2 to 6.2.4.4.

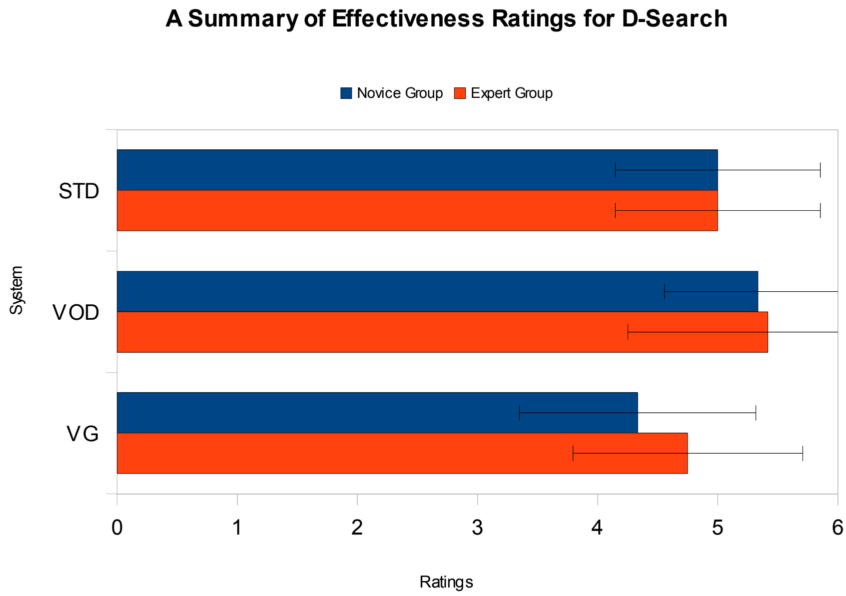


Figure 6.16 A Summary of User Effectiveness Ratings for D-Search

6.2.4.2 Efficiency Ratings

For the question that "I got results quickly with D-Search (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in the first table in Table 6.43 and the relevant statistical results are shown in the second table where t-test was used. The two groups reported same mean value (4.9167) which indicated a quite positive response to the efficiency of D-Search. There is no significant difference between the two groups (P = 1.000). Since our results were produced and analysed on the basis of groups, it indicates that the positive ratings were system-dependent rather than group-dependent.

Group	Size	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	0	1	3	4	4	0.00	0.00	8.33	33.33	33.33	33.33
Expert	12	0	0	1	3	4	4	0.00	0.00	8.33	25.00	33.33	33.33

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	4.9167	0.9962	0.2876	0.000	1.000
Expert	4.9167	0.9962	0.2876		

Table 6.43 Independent User Efficiency Ratings for D-Search (P > 0.05)

For the question that "I got results more quickly with D-Search than the Open Directory (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in Table 6.44. The two groups reported similar mean value (> 5.0000) which indicated very positive response to the efficiency of D-Search against the Open Directory. Although the expert group gave slightly higher average ratings, there is no significant difference between the two groups (P = 0.430).

Group	Size	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	0	1	0	7	4	0.00	0.00	8.33	0.00	58.33	33.33
Expert	12	0	1	0	0	2	9	0.00	8.33	0.00	0.00	16.67	75.00

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	5.1667	0.8349	0.3371	0.804	0.430
Expert	5.5000	1.1678	0.2410		

Table 6.44 User Efficiency Ratings of D-Search Compared to the Open Directory (P > 0.05)

For the question that "I got results more quickly with D-Search than Google (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in Table 6.45. The two groups reported similar mean value between 4.0000 and 5.0000 which indicated slightly positive response to the efficiency of D-Search against Google. Similarly, although the expert group gave slightly higher average ratings, there is no significant difference between the two groups (P = 0.881).

Group	Size	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	1	3	3	3	2	0.00	8.33	25.00	25.00	25.00	16.67
Expert	12	0	2	1	4	2	2	0.00	16.67	8.33	33.33	16.67	16.67

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	4.1667	1.2673	0.3658	0.152	0.881
Expert	4.2500	1.4222	0.4106		

Table 6.45 User Efficiency Ratings of D-Search Compared to Google (P > 0.05)

Figure 6.17 presents an overview of the mean and SD from Table 6.43 – 6.45 where showed overall positive ratings and rating trends by both groups. The expert group always gave more positive ratings than the novice group. Both groups tended to give higher efficiency ratings when comparing D-Search and the Open Directory whereas they tended to lower the ratings when comparing to Google.

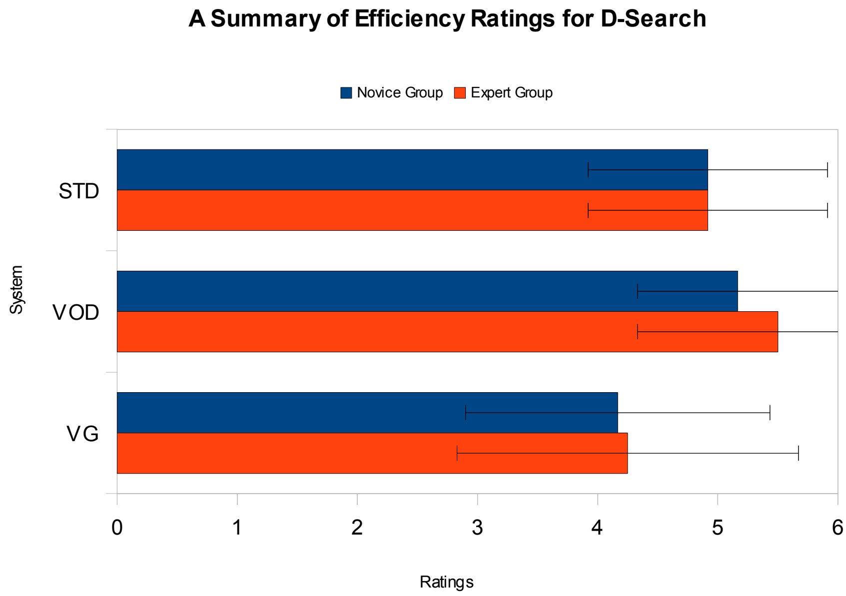


Figure 6.17 A Summary of User Efficiency Ratings for D-Search

6.2.4.3 Overall Satisfaction Ratings

For the question that "I am satisfied with D-Search for the tasks (1 strongly disagree – 6 strongly agree)", the detailed user ratings given by the two groups are shown in the first table in Table 6.46 and the relevant statistical results are shown in the second table where t-test was used. The novice user group reported slightly positive ratings (4.5000) while the expert group reported highly positive ratings (5.0833) for the satisfaction of D-Search. There is no significant difference between the two groups (P = 0.065) which indicated that the positive ratings were system-dependent rather than group-dependent.

Group	Size	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	0	1	5	5	1	0.00	0.00	8.33	41.67	41.67	8.33
Expert	12	0	0	0	2	7	3	0.00	0.00	0.00	16.67	58.33	25.00

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	4.5000	0.7977	0.2303	1.941	0.065
Expert	5.0833	0.6686	0.1930		

Table 6.46 Independent User Satisfaction Ratings for D-Search (P > 0.05)

For the question that "I am more satisfied with D-Search than the Open Directory for the tasks (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in Table 6.47. The two groups reported similar mean value (> 5.0000) which indicated very positive response to the satisfaction of D-Search against the Open Directory. Although the expert group gave slightly higher average ratings, there is no significant difference between the two groups (P = 0.213).

Group	**Size**	**Size**						**Percentage (%)**					
		1	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**
Novice	12	0	0	1	2	4	5	0.00	0.00	8.33	16.67	33.33	41.67
Expert	12	0	0	0	1	4	7	0.00	0.00	0.00	8.33	33.33	58.33

Group	**Mean**	**Std. Deviation**	**Std. Error Mean**	**t**	**P**
Novice	5.0833	0.6742	0.2876	1.200	0.213
Expert	5.5000	0.9962	0.1946		

Table 6.47 User Satisfaction Ratings of D-Search Compared to the Open Directory (P > 0.05)

For the question that "I am more satisfied with D-Search than Google for the tasks (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in Table 6.48. The two groups reported same mean value (4.8333) close to 5.0000 which indicated quite positive response to the satisfaction of D-Search against Google. There is no significant difference between the two groups (P = 1.000).

Group	**Size**	**Size**						**Percentage (%)**					
		1	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**
Novice	12	0	0	1	3	5	3	0.00	0.00	8.33	25.00	41.67	25.00
Expert	12	0	0	1	4	3	4	0.00	0.00	25.00	33.33	25.00	33.33

Group	**Mean**	**Std. Deviation**	**Std. Error Mean**	**t**	**P**
Novice	4.8333	0.9374	0.2706	0.000	1.000
Expert	4.8333	1.0299	0.2973		

Table 6.48 User Satisfaction Ratings of D-Search Compared to Google (P > 0.05)

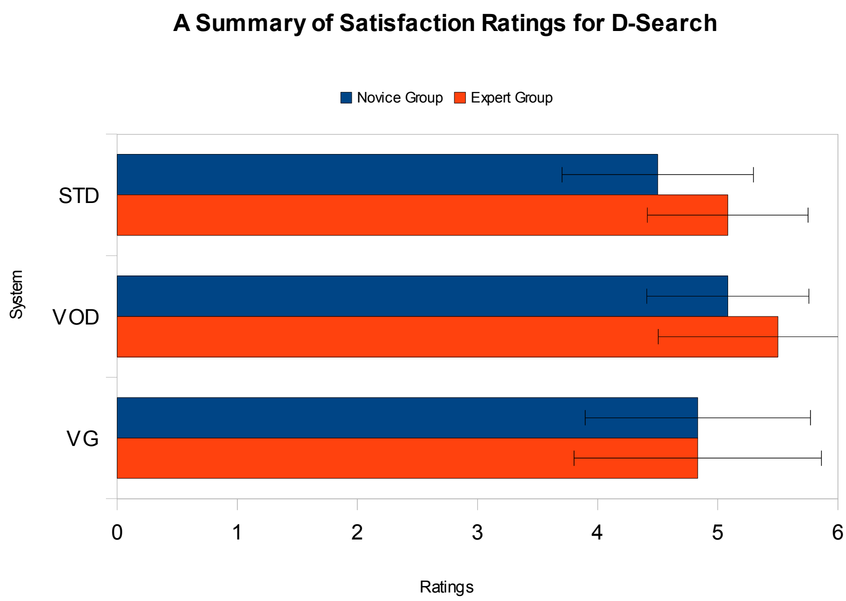


Figure 6.18 A Summary of User Satisfaction Ratings for D-Search

Figure 6.18 presents an overview of the mean and SD from Table 6.46 – 6.48 where showed overall positive ratings and rating trends by both groups. The expert group always gave more positive ratings than the novice group. Both groups tended to give higher efficiency ratings when comparing D-Search and the Open Directory whereas they tended to lower the ratings when comparing to Google.

6.2.4.4 Search Helpfulness Ratings

For the question that "D-Search is helpful in locating websites within a Web directory (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in the first table in Table 6.49 and the relevant statistical results are shown in the second table where t-test was used. Both groups gave highly positive ratings over 5.000 for the satisfaction of D-Search. There is no significant difference between the two groups (P = 0.519) which indicated that the positive ratings were system-dependent rather than group-dependent.

Group	Size	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	1	0	0	5	6	0.00	0.00	0.00	0.00	50.00	50.00
Expert	12	0	0	0	1	4	7	0.00	0.00	0.00	8.33	33.33	66.67

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	5.2500	1.1382	0.3286	0.655	0.519
Expert	5.5000	0.6742	0.1946		

Table 6.49 Independent User Search Helpfulness Ratings of D-Search (P > 0.05)

For the question that "D-Search is more helpful than the original Open Directory in locating websites within a Web directory (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in Table 6.50. The two groups reported similar highly positive mean value (> 5.5000) which indicated a very positive acknowledgement to the search helpfulness of D-Search against the Open Directory. Although the expert group gave slightly higher average ratings, there is no significant difference between the two groups (P = 0.430).

Group	Size	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	0	0	0	6	6	0.00	0.00	0.00	0.00	50.00	50.00
Expert	12	0	0	0	0	4	8	0.00	0.00	0.00	0.00	33.33	66.67

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	5.5000	0.5222	0.1508	0.804	0.430
Expert	5.6667	0.4924	0.1421		

Table 6.50 User Search Helpfulness Ratings of D-Search Compared to The Open Directory (P > 0.05)

For the question that "D-Search is more helpful than Google in locating websites (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in Table 6.51. The two groups reported similar highly positive mean value (>5.0000) which indicated quite positive response to using D-Search to search against Google. There is no significant difference between the two groups (P = 0.223).

Group	Size	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	1	0	4	2	6	0.00	0.00	0.00	33.33	16.67	50.00
Expert	12	0	0	0	1	3	8	0.00	0.00	0.00	8.33	25.00	66.67

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	5.1667	0.9374	0.2706	1.254	0.223
Expert	5.5833	0.6686	0.1930		

Table 6.51 User Search Helpfulness Ratings of D-Search Compared to Google (P > 0.05)

Figure 6.19 presents an overview of the mean and SD from Table 6.49 – 6.51 where showed very high positive ratings from both groups for the search helpfulness of D-Search no matter if comparing to other systems or not.

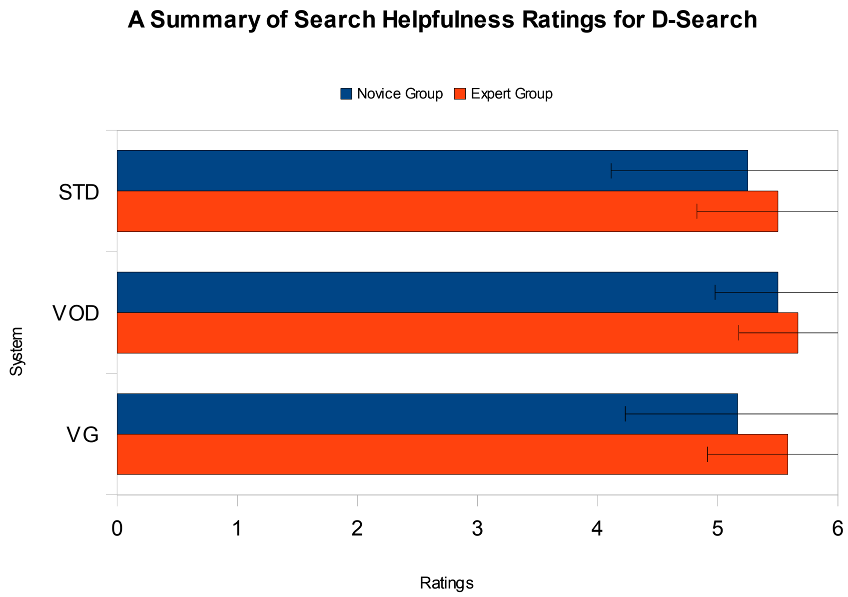


Figure 6.19 A Summary of Search Helpfulness Ratings for D-Search

In summary, users gave over 4 ratings in all assessments of D-Search. Moreover, their standalone ratings are generally similar to the comparative ratings to Google which indicates similar performance of D-Search and Google. However, since their standalone ratings are generally lower to the comparative ratings to the Open Directory. This indicates D-Search had better performance than the Open Directory.

6.2.5 Subjective Feedbacks

The subjective feedback about users' experience with D-Search was collected by using an open question asking participants to give their opinion about D-Search. In this part, both positive messages (talking points) and negative messages were raised from their answers. In general, positive responses to D-Search are focused into the following three areas.

- Participants acknowledged the conceptual match between D-Search and the Open Directory in terms of in-directory search. They agreed that D-Search presented more practical search need than conventional Web search models. That is, to accelerate users' navigation in looking for specific categories in a Web directory.
- Participants stressed the search advantages and convenience of exemplar-based D-Search compared to normal keyword-based Web search engines in terms of searching through Web directory content.
- Participants said the easy-to-use D-Search and accurate search results brought them more confidence in using Web directories and they considered D-Search as a necessary facility for any general Web directory.

Some negative responses to D-Search were also raised at the same time which were mainly concerned on two side of D-Search.
For interaction side,

- Some participants said the current input syntax for D-Search was somewhat strict as it only supported the standard URL format like http://www.foo.com/" instead using common URL inputs such as "www.foo.com" or "foo.com". Compared to the natural word input for Google and other Web search engines, this is not convenient.
- Several participants said to initialise an exemplar was sometimes not easy for various reasons such as misspelling an address or misremembering an address.

For technical side, few participants reported that sometimes D-Search was unable to find relevant categories because the address was not in the directory.

6.3 Results of D-Persona

D-Persona is an implementation of the content-based personalisation model in our unified framework. It delivers a tailored directory to a user's interests based on their browsing profile. We only collected preference data from the open user task for measuring the outcome of it. Thus, results of D-Persona are reported in two measurements, user feedback ratings and subjective feedbacks.

6.3.1 User Feedback Ratings

In the same way used for capturing participants' feedback for D-Search, we used a one-to-six rating system to capture user feedbacks on D-Persona in a quantitative data form. We considered the user ratings of the effectiveness, efficiency, overall satisfaction and browsing helpfulness. Note results are only produced in group-oriented views.

6.3.1.1 Effectiveness Ratings

For the question that "I found results successfully with D-Persona (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups for D-Persona are shown in the first table in Table 6.52 and the relevant statistical results are shown in the second table where t-test was used. Both groups gave highly positive ratings over 5.000 for the effectiveness of D-Persona. There is no significant difference between the two groups (P = 0.223) although the expert group reported higher ratings than the novice group. This indicates that the positive ratings were measurement-dependent rather than group-dependent.

Group	Size	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	1	0	2	4	5	0.00	8.33	0.00	16.67	33.33	41.67
Expert	12	0	0	0	1	4	7	0.00	0.00	0.00	8.33	33.33	58.33

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	5.0000	1.2061	0.3482	1.254	0.223

Expert	5.5000	0.6742	0.1946		

Table 6.52 Independent Effectiveness Ratings of D-Persona (P > 0.05)

6.3.1.2 Efficiency Ratings

For the question that "I got results quickly with D-Persona (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups for D-Persona are shown in the first table in Table 6.53 and the relevant statistical results are shown in the second table where t-test was used. Both groups gave highly positive ratings at around 5.000 for the efficiency of D-Persona. It is interesting to see that the novice group gave higher ratings than the expert group. However, there is no significant difference between the two groups (P = 0.738).

Group	**Size**	**Size**						**Percentage (%)**					
		1	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**
Novice	12	0	0	1	1	6	4	0.00	0.00	8.33	8.33	50	33.33
Expert	12	1	0	0	2	4	5	8.33	0.00	0.00	16.67	33.33	41.67

Group	**Mean**	**Std. Deviation**	**Std. Error Mean**	**t**	**P**
Novice	5.0833	0.9003	0.2599	-0.339	0.738
Expert	4.9167	1.4439	0.4167		

Table 6.53 Independent Efficiency Ratings of D-Persona (P > 0.05)

6.3.1.3 Overall Satisfaction Ratings

For the question that "I am satisfied with D-Persona (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in the first table in Table 6.54 and the relevant statistical results are shown in the second table where t-test was used. Both groups gave positive ratings (>4.5000) for the efficiency of D-Persona but similar to efficiency ratings, the novice group gave higher ratings than the expert group. However, there is no significant difference between the two groups (P = 0.328).

Group	**Size**	**Size**						**Percentage (%)**					
		1	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**
Novice	12	0	0	1	2	5	4	0.00	0.00	8.33	16.67	41.67	33.33
Expert	12	0	1	0	4	5	2	0.00	8.33	0.00	33.33	41.67	16.67

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	5.0000	0.9535	0.2752	-1.000	0.328
Expert	4.5833	1.0836	0.3128		

Table 6.54 Independent Satisfaction Ratings of D-Persona (P > 0.05)

6.3.1.4 Browsing Helpfulness Ratings

For the question that "D-Persona is helpful when using the Open Directory (1 strongly disagree – 6 strongly agree)", the detailed ratings given by the two groups are shown in the first table in Table 6.55 and the relevant statistical results are shown in the second table where t-test was used. Both groups gave positive ratings at around 5.0000 for the browsing helpfulness of D-Persona. There is no significant difference between the two groups (P = 0.865).

Group	Size	Size						Percentage (%)					
		1	2	3	4	5	6	1	2	3	4	5	6
Novice	12	0	1	0	1	7	3	0.00	8.33	0.00	8.33	58.33	25.00
Expert	12	0	1	0	3	2	6	0.00	8.33	0.00	25.00	16.67	50.00

Group	Mean	Std. Deviation	Std. Error Mean	t	P
Novice	4.9167	1.0836	0.3128	0.172	0.865
Expert	5.0000	1.2792	0.3693		

Table 6.55 Independent Browsing Helpfulness Ratings of D-Persona (P > 0.05)

In summary, users gave nearly 5 ratings in all assessments which indicates highly positive responses to D-Persona.

6.3.2 Subjective Feedbacks

The subjective feedback about users' experience with D-Persona was collected by using a similar open question to D-Search to ask participants to give their opinion. We found both positive messages (talking points) and negative messages from their answers. In general, good comments focused on the overall results quality of D-Persona in terms of users' profile. Some also mentioned the practical use and importance of D-Persona in helping users navigate through the Open Directory. As we saw in the feedback for D-Search, adverse comments for D-Persona concentrated on the technical side. For example, some participants thought the recall of categories was sometimes poor because some websites showing in user profiles were not collected by the directory. This issue was anticipated as we have adopted the same content match mechanism for implementing D-Search and D-Persona.

6.4 Discussion

The main purpose of our comparative user test is to assess the usability of our unified framework in terms of the separate implementations corresponding to its two sub models. That is, D-Search for the redefined search model and D-Persona for the individual content-based personalisation model. Thus, the results of D-Search and D-Persona are discussed individually in terms of the usability metrics we used in the measurement.

6.4.1 Judgement of Hypotheses for D-Search

6.4.1.1 Judgement of H1

H1 states that "a user will complete search tasks more quickly with D-Search than with the original Open Directory and Google" (Chapter 5 Section 5.4.1). In our experiment, it equally means "a user, *either novice or expert*, will complete search tasks, *either simple or complex* more quickly with D-Search than with the original Open Directory and Google". This was analysed by using the task completion time as performance data measurement and user feedback ratings as preference data measurement. Table 6.56 shows the needed p-value in the task completion time analysis for proving the hypothesis.

H1	Task	User		System
		Novice	Expert	DS
	Simple	P < 0.05	P > 0.05	P > 0.05
	Complex	P > 0.05	P > 0.05	P > 0.05

Table 6.56 The P-value in Task Completion Time for H1

Before considering the actual value (i.e., task completion time) on the comparative systems, the first step to prove H1 is true requires significant system difference ($p < 0.05$) in the "User" column and insignificant group difference ($p > 0.05$) in the "System" column of the table. However, except for the significant system difference ($p < 0.05$) reported in the novice user group for simple tasks, the insignificant system difference ($p > 0.05$) was found in the rest which indicates H1 has been rejected although the factual data showed clear difference. Since one-way ANOVA is a parametric analysis, this outcome is more likely a side-effect of the small sample size/population of each user group we had in the experiment ($n = 8$) as the size is too small to gain the sufficient power to detect any significant difference among the population. Possible solutions include running either a non-parametric analysis like Wilcoxon signed-rank test for twice comparisons (D-Search vs. Google and D-Search vs, the Open Directory) or a factual data analysis. Here we take the second approach for the discussion. We found both groups completed simple tasks on D-Search more quickly than Google and the Open Directory while the novice group spent less time on Google than D-Search then the Open Directory and the expert group spent less time on the Open Directory than D-Search then Google. For the novice group, consider the average time they spent within the two systems (Google: 58.25 seconds, D-Search: 94.31 seconds), it indicates they took significantly loner time to check the relevancy of results on D-Search than Google. This is due to as a natural "disadvantage" of any Web directory when a user is searching a specific product (e.g., 512MB Sandisk SD card) with less search experience as the directory cannot tell whether websites in a matched category sell the product. In comparison, Google can do this directly with its results if the query contains such keywords. Thus, unlike the expert group who were able to "pick" relevant websites quickly (Google: 37.29 seconds, D-Search: 35.15 seconds) based on the description of these websites, the novice user group had to spent more time on trials. From this point of view, it is better to consider the user feedback ratings to understand H1 from a general perspective. For the expert group, consider the task success rate in regarding to understand the task completion time, we found most users in the group failed on the Open Directory (25%) as they gave up very early. Since we have encouraged them to run the tasks if we did not see adequate user efforts in the test, the success rate suggests that the task completion time on the Open Directory should not be used to compare to more successful systems like D-Search (100%) and Google (87.5%). From this point of view, we could prove the H1 in the situation when successful completion of tasks is required.

Table 6.57 shows the p-value of user feedback ratings in the efficiency analysis.

	Independent	**Compared to the Open Directory**	**Compared to Google**
H1	$P > 0.05$	$P > 0.05$	$P > 0.05$

Table 6.57 The P-value in User Feedback Ratings for H1

With an average of 4.92, both groups gave highly positive ratings in the independent testing statement for the efficiency of D-Search where no rating difference was found between the two group ($p > 0.05$). The ratings jumped to 5.17 and 5.5 respectively when compared to the Open Directory ($p > 0.05$) and dropped a little to 4.17 and 4.25 respectively when compared to Google ($p > 0.05$). Consider the results in performance measurement, this is more likely to reflect the users' feeling when a successful task completion is needed to achieve.

Therefore, according to the above discussion, we can prove that, to any user, “they will complete search tasks more quickly with D-Search than with the original Open Directory and Google in the synthetic consideration (e.g., effectiveness in completing tasks)”.

6.4.1.2 Judgement of H2

H2 states that “a user will complete search tasks more successfully with D-Search than with the original Open Directory and Google" (Chapter 5 Section 5.4.1). In our experiment, it equally means "a user, *either novice or expert*, will complete search tasks, *either simple or complex* more successfully with D-Search than with the original Open Directory and Google". This was analysed by using the normalised success rate as performance measurement and user feedback ratings as preference measurement.

Table 6.58 shows the needed p-value in the normalised success rate analysis for proving H2.

H2	Task	User		System
		Novice	Expert	DS
	Simple	P < 0.05	P < 0.05	P > 0.05
	Complex	P < 0.05	P < 0.05	P > 0.05

Table 6.58 The P-value in Normalised Success Rate for H2

In the "User" column, significant system difference ($p < 0.05$) was found for both groups performing simple and complex tasks which indicates the acceptance of H2 (systems show difference) and rejection of its h0 (systems show no difference). Moreover, both groups reported the highest normalised success rate on D-Search in complex tasks where 87.5% was for the novice group and 91.7% was for the expert group. Since no group difference was found on D-Search in the "System" column, it can prove H2 in the situation when performing complex tasks. However, a slightly lower normalised success rate on D-Search (85.4%) than Google (87.5%) was found with the novice users while same success rate on Google (87.5%) and D-Search (87.5%) was found with the expert users although both groups achieved 100% task success rate. Consider they spent significant shorter time on D-Search (novice: 20.79 seconds, expert: 11.48 seconds) than Google (novice: 178.74 seconds, expert: 119.66 seconds), it could prove H2 in the situation when performing simple tasks with the efficiency requirement.

Table 6.59 shows the p-value of user feedback ratings in the effectiveness analysis.

	Independent	Compared to the Open Directory	Compared to Google
H2	P > 0.05	P > 0.05	P > 0.05

Table 6.59 The P-value in User Feedback Ratings for H2

Both groups showed highly favourable responses to the effectiveness of D-Search by giving an average of 5 where no significant group rating difference is found ($p > 0.05$). This indicates the ratings have no relation to the group diversity (e.g., user expertise and experience etc.) but are only related to the context of the testing statements. Similar results (good ratings and $p > 0.05$) were also found in the comparative ratings to the Open Directory and Google. In addition, both groups tended to raise their ratings in comparison to the Open Directory but to lower their ratings in comparison to Google. This is also consistent with our findings in performance measurement where they showed poor normalised success rate on the Open Directory but still good success rate on Google.

According to the above discussion for the results of performance measurement and preference measurement, we can prove that, to any user, “they will complete tasks more successfully with D-Search than with the original Open Directory and Google in the synthetic consideration (e.g., the efficiency in completing simple tasks)”.

6.4.1.3 Judgement of H3

H3 states that “a user will be more satisfied in completing tasks with the D-Search than the original Open Directory or Google” (Chapter 5 Section 5.4.1). This was measured by using preference data (i.e., user feedback ratings) with t-test. Table 6.60 shows the P-value of t-test for the independent satisfaction ratings and two comparative satisfaction ratings.

	Independent	**Compared to the Open Directory**	**Compared to Google**
H3	$P > 0.05$	$P > 0.05$	$P > 0.05$

Table 6.60 The P-value in User Feedback Ratings for H3

Good ratings were received from both groups in the independent test and the expert group (avg. 5.08) responded more positively than the novice group (avg. 4.5). However, since $p > 0.05$ in the t-test, the rating difference between two groups is not significant. Similar results (e.g., good ratings, $P>0.05$) were also found in the comparative rating test to the Open Directory and Google. It should also be noted that users in either group tended to give similar ratings in the independent test and the comparative test to Google while they tended to give better ratings in the comparative test to the Open Directory than in the independent test. These results are consistent with the results of task completion time and normalised success rate in preference measurement. Thus, we can prove H3 that, to any user, "they will be more satisfied in completing tasks with the D-Search than the original Open Directory or Google".

6.4.1.4 Judgement of H4

H4 states that "a user will find using D-Search to search the content of the Open Directory is more helpful than its original search engine or Google in terms of their intentions" (Chapter 5 Section 5.4.1). Similar to H3, this was analysed by using user feedback ratings with t-test. Table 6.61 shows the P-value of t-test for the independent helpfulness ratings and two comparative helpfulness ratings.

	Independent	Compared to the Open Directory	Compared to Google
H4	$P > 0.05$	$P > 0.05$	$P > 0.05$

Table 6.61 The P-value in User Feedback Ratings for H4

Both groups gave highly positive ratings (avg. > 5) in either independent test or comparative tests where the rating difference between two groups is not significant ($p > 0.05$). This is also consistent with the findings in our performance measurement. Thus, we can prove H4 that "to any user, they will find using D-Search to search the content of the Open Directory is more helpful than its original search engine or Google in terms of their intentions".

6.4.2 Judgement of Hypotheses for D-Persona

Unlike D-Search, all four hypotheses for D-Persona were analysed by preference measurement with t-test as a user's personalised results are not comparable to others in terms of individual content-based personalisation. Table 6.62 shows the p-value of t-test for all user feedback ratings related to the hypotheses.

	Independent	Compared to the Open Directory	Compared to Google
H1	P > 0.05	P > 0.05	P > 0.05
H2	P > 0.05	P > 0.05	P > 0.05
H3	P > 0.05	P > 0.05	P > 0.05
H4	P > 0.05	P > 0.05	P > 0.05

Table 6.62 The P-value in User Feedback Ratings for D-Persona

H1 states that “a user will find his interested content successfully with D- Persona” (Chapter 5 Section 5.4.2). Both groups showed highly favourable responses to the effectiveness of D-Persona by giving an average of 5 and no significant rating difference is found between the two groups ($P>0.05$). This indicates the ratings have no relation to the group diversity (e.g., user expertise and experience etc.) but are only related to the context of the testing statement. Thus, we can prove H1 that, “to any user, they will find their interested content successfully with D-Persona”.

H2 states that “a user will find his interested content quickly with D-Persona”. Both groups responded highly positively to the efficiency of D-Persona by giving 5.08 and 4.92 respectively. Although the novice group gave slightly higher ratings than the expert group, there is no significant difference between the two groups ($p > 0.05$). Thus, we could prove H2 that, to any user, “they will find their interested content quickly with D-Persona”.

Similarly to H1 and H2, as highly positive ratings were found in both groups and $p > 0.05$ was also found in all related testing statements for H3 and H4, we can prove H3 that, to any user, “they will be more satisfied with their navigation quality of D-Persona” and H4 that, “they will find D-Persona improves their judgement in the helpfulness of using the Open Directory".

Therefore, we conclude that our hypotheses of D-Search and D-Persona were completely examined and proved so that our primary purpose of the user study was fulfilled.

6.4.3 Implications and Others

The user pathway data were introduced as a supplementary performance measurement to task completion time and normalised success rate for D-Search. Table 6.63 shows the p-value in the measurement of user pathway data where the top part are from group-oriented results and the bottom part are from system-oriented results.

Task	User		System
	Novice	Expert	DS
Simple	P < 0.05	P < 0.05	P > 0.05
Complex	P < 0.05	P < 0.05	P > 0.05

Table 6.63 The P-value in User Pathway Data for D-Search

Both groups reported significant difference ($p < 0.05$) in using queries for completing tasks on the comparative systems. The novice group generated an average of 1.375 queries on D-Search and 1.625 queries on Google for simple tasks and 1.125 and 2.375 queries for complex tasks respectively. Similarly, the expert group generated an average of 1.25 queries on D-Search and 1.75 queries on Google for simple tasks and 1.25 and 1.5 queries for complex tasks respectively. The quantity change of queries on D-Search is less significant than Google when the task complexity rises. This indicates D-Search is more stable than Google to any user in terms of the use of search queries. On the other hand, it also means that D-Search is more suitable to novice users than expert users. This can be also found from the normalised success rate where the rate dropped significantly with the novice group when the task complexity grew. In addition, the similar number of query usage on D-Search and Google also shows that our predefined tasks have balanced the different search strength of Web directories and search engines.

Moreover, although users reported that sometimes it was not easy to initialise a query on D-Search, the success rate and the small number of queries show that initialising exemplar on D-Search is as easy as initialising a query on Google.

We also found that although both groups launched similar number of queries for simple tasks on D-Search and Google, the task completion time on D-Search is much shorter than Google. This indicates that D-Search is more suitable than Google for the Open Directory as the simple tasks represented the fundamental search needs in the directory which also proves that document search like what a Web search engine does is not suitable for using in Web directories. Since the completion time are similar in complex tasks, it implies that even though Web directories are not specifically designed for some complex search needs, D-Search would help users improve their search efficiency for these tasks.

6.5 Summary

The results of the comparative user study and relevant discussions have been introduced in this chapter. The study demonstrates the strengths of our unified framework in terms of its two separate implementations called D-Search and D-Persona. Their outcomes also prove that the framework can be used as a user-centred solution for improving the overall user navigational experience in Web directories.

Chapter 7 Conclusion

- *Will there be a usability approach to help users improve their navigational experience in Web directories?*
- *Yes, we have proposed a unified framework and tested it with significant experimental results in improving the user experience of Web directories.*

7.1 Conclusion

The key characteristic of this thesis is its role of defining the classificatory rigidity of Web directories on theoretical grounds, verifying it through usability inspections and assessing the proposed framework for improving it from a user study.

Two classification schemes have been reviewed and compared for their suitability as a classification scheme for generic Web directories. Hierarchical classifications are currently used in all large generic Web directories, including Yahoo! Directory and Open Directory, for organising their massive website collections while faceted classifications are commonly adopted for managing the product collections of e-commerce websites. We discovered that rigidity, which derives from the strict structural requirement of hierarchies, lies in almost every principle of establishment from viewing a knowledge domain, defining class inclusions, divisions and relationships to making cross-references. We also identified that the poor user experience of large hierarchical classification systems is rooted in rigidity. Compared to hierarchies which link entities and classes in a general – specific relationship with a consistent view, we noticed that facets are free from the constraints described in hierarchies. This allows faceted classifications to produce far less rigid representations. However, such freedom makes faceted classifications unsuitable for organising large generic Web directories as they are unable to represent the need for Web directories – to guide and help user navigate the Web in certain point of views. In the light that this purpose can be only achieved with hierarchical classifications, we therefore concluded that hierarchical classifications are still the first choice for Web directories and hence suggested to focus on their representations (user interface) for possible improvements.

Subsequently, we adopted Ontological Sketch Modelling (OSM), which is a conceptual-based usability inspection method, to conduct a user study for understanding how rigidity affects user navigation on the representational level of Web directories with hierarchical classifications. The misfits between users and Web directories reaffirmed the case that rigidity is the main cause for generating navigational difficulties. That is, the classificatory and cross-referencing standards used by a Web directory do not always match a user's understanding of them. We then decided to look into Web personalisation to eliminate the rigidity of hierarchical Web directories. Web personalisation is a research field that asserts that user experience of an information system relies on the content of the system. In other words, a user will always have problems during navigation as long as he needs to spend efforts on distinguishing a genuine piece of information he is interested in. On the other hand, we also discovered that the current search model used in Web directories (i.e., a simplified and localised Web search model) could not meet the true user demand for searching Web directories. That is, like searching in a traditional classification system has the goal of finding out the location of a publication, searching in a Web directory is directed at finding a category which represents a user's interest. For this reason, we proposed a new search model which offers good category locating ability as a more suitable for meeting the search demand for Web directories.

An individual content-based personalisation model and a redefined search model were unified in one name-space matching mechanism under a proposed framework for user experience improvement. The idea was to allow users to use website exemplars (i.e., some expected results) to retrieve corresponding categories in the directory. Theoretically, a website is the best descriptor of its category because not only does the entity inherit all attributes of the class, it is also considered as a proper representative for the class. Thus, a website exemplar (name) reflects the only true intention (space) of the user so that guarantees the accuracy of search. This framework was assessed in a comparative user study. Results showed that both implementations of the two models achieved satisfying performance in all key measurements including efficiency, effectiveness, user satisfaction and feedbacks, especially for the search model's performance in comparison to Google and the Open Directory. Therefore, we concluded that the unified framework can be used as a user-centric solution to improve the user experience in Web directories.

7.2 Future Work

This approach of the unified framework is a branch “grafted” onto existing research which is still more experimental than practical although the user study has demonstrated its significance. Thus, some improvements would be needed in the future. First, an explicit profiling technique was applied for learning a user's interests in the past with some user-level approvals in the personalisation model. The disadvantage is that it requires more user effort during the process and the profile accuracy would be reduced more quickly over time than implicit user profiling. In the future, implicit profiling techniques can be introduced to build and adapt user profiles automatically. Second, the name-space matching mechanism used by both personalisation and only considered individual's demand so it would easily cause a poor recall when a user's website exemplar is not listed in the Web directory. A possible solution is to introduce collaborative content-match mechanisms to enable user-to-user recommendations. In addition, for the redefined search, prompts including auto-correction and auto-complete could be considered in the future improvements.

7.3 Contributions

The range of contributions made by this research includes:

- A proper definition of rigidity in hierarchies;
- A theoretical validation of rigidity as the main cause of poor user navigation in Web directories;
- A general guideline for determining suitable classification schemes for general Web directories.
- A clarified and redefined user search model of Web directories;
- A unified framework featuring enhanced browsing and searching experience in Web directories.

In conclusion, taking Web directories as the main entry point, the thesis presented a new research direction to improve user experience without losing the organisational strengths of classifications at the same time, which could be expanded to any information classifications with hierarchical representations.

Appendix: The Forms and Data of the Experiment

Part 1 Forms Used in the Experiment

1.1 User Background Survey and Level Test

User Background	
Please rate your use frequency of the Internet.	Heavy (4+ hrs per day)
	Regular (1 – 4 hrs per day)
	Light (<1hr per day)
Please tell us your activities on the Internet (multiple choices).	Reading (news, blogs, forum/BBS)
	Communicating (Emails, IM chatting,voice chatting etc.)
	Gaming
	Shopping (incl. selling and renting)
	Streaming (movie, music, radio, tv)
	Downloading
	Researching
Please list your interested topics (10 in maximum) based on your activities. For example, if shopping, list the broad product category you are interested (write your answer below): *<write your answer here>*	
Please rate your use frequency of Web search engines.	Quite often (whenever having a search demand, I would use it)
	Only several times
	Heard of it but never used before
	Never heard of it
Please tell us how often you experience difficulties in using a search engine.	Never
	Only several times
	Half, half
	Quite often
	Almost every time
	Never used a search engine before
If you have used search engines before, please state the worst experience you have had with a search engine (write your answer below): *<write your answer here>*	
Please rate your use frequency of Web directories.	Quite often
	Only several times
	Heard of it but never used before

	Never heard of it
Please tell us how often you experience difficulties in using a Web directory.	Never
	Only several times
	Half, half
	Quite often
	Almost every time
	Never used
If you have used Web directories before, please state the worst experience you have had with a Web directory (write your answer below): *<write your answer here>*	
User Level Test	
Test 1 (Web Directories)	**Test 2 (Search Engines)**
First look at the homepage of Open Directory, can you tell which category could be the most appropriate category for amazon.co.uk? Category name: _____________ Based on your answer, can you find it? Task completed in ____________minutes.	A recent article on the BBC website revealed that Britain conducted a similar survey to Kinsey report in 1940s. Since the source of the report was not mentioned, can you identify the survey (name, summary) by using Google? Task completed in ____________minutes.

1.2 User Feedback Ratings Form

1.2.1 Form for D-Search

User Ratings	
Please rate the following statements for D-Search. (1 strongly disagree – 6 strongly agree)	
I found results successfully with D-Search.	1 2 3 4 5 6
I got results quickly with D-Search.	1 2 3 4 5 6
I am satisfied with D-Search for the tasks.	1 2 3 4 5 6
D-Search is helpful in locating websites within a Web directory.	1 2 3 4 5 6
Please rate the following statements for D-Search in comparison to Google. (1 strongly disagree – 6 strongly agree)	
I found results more successfully with D-Search than Google.	1 2 3 4 5 6
I got results more quickly with D-Search than Google.	1 2 3 4 5 6
I'm more satisfied with D-Search than Google for the tasks.	1 2 3 4 5 6
D-Search is more helpful than Google in locating websites within a Web directory.	1 2 3 4 5 6
Please rate the following statements for D-Search in comparison of the Open Directory. (1 strongly disagree – 6 strongly agree)	
I found results more successfully with D-Search than the Open Directory.	1 2 3 4 5 6

I got results more quickly with D-Search than the Open Directory.	1 2 3 4 5 6
I'm more satisfied with D-Search than the Open Directory for the tasks.	1 2 3 4 5 6
D-Search is more helpful than the Open Directory in locating websites.	1 2 3 4 5 6
Please leave any other comments you would like to add (for example, how do you think the personalisation can improve your navigation in a Web directory, what do you feel about using this function in terms of its advantages and disadvantages, are you enjoyed with D-Persona etc.)	

1.2.2 Form for D-Persona

List of questions (Questionnaire)
1. Please rate for the following topics (1 – strongly disagree, 6 – strongly agree): 1.1 I found results successfully with D-Persona. (1 2 3 4 5 6) 1.2 I got results quickly with D-Persona. (1 2 3 4 5 6) 1.3 I am satisfied with D-Persona. (1 2 3 4 5 6) 1.4 D-Persona is helpful when using the Open Directory. (1 2 3 4 5 6) **2. Please leave any other comments you would like to add (for example, how do you think the personalisation can improve your navigation in a Web directory, what do you feel about using this function in terms of its advantages and disadvantages, are you enjoyed with D-Persona etc.)**

1.3 Data Recording Form

User login:	**Task (circle it): Simple 1 2 3 Complex 1 2 3**	
User queries (continue to write on the reverse side if needed)		Time spent:
		Time spent:
		Time spent:
		Time spent:
		Time spent:
		Time spent:
		Time spent:
		Time spent:
		Time spent:
		Time spent:

Part 2: Experimental Data of D-Search

2.1 Novice User Group Results for Simple Tasks

TSK represents for the type of task the user performed where s(1/2) indicates simple task 1/2.

PLT represents for the platform where the user performed the task, where DS represents for D-Search, G represents for Google and OD represents for the Open Directory;
T O/S represents for the average completion time the user took in query sessions;

O/S Fails represents for the number of queries failed by the user;

T I/S represents for the average completion time the user spent in external links;

QRY represents for the number of queries used by the user;

TSK Fails represents for the number of failed tasks;

TCT represents for the total task completion time as a sum of AVG O/S and AVG.

USR	TSK	PLT	N O/S	T O/S	O/S Fails	T I/S	QRY	TSK Fails	TCT
iyj1	s1	DS	3	5.53	2	0	3	0	5.53
jib1	s1	DS	1	12.5	0	0	1	0	12.5
mrs1	s1	DS	2	8.3	1	0	2	0	8.3
uss1	s1	DS	1	56.6	0	0	1	0	56.6
inm1	s2	DS	1	36.1	0	0	1	0	36.1
lir1	s2	DS	1	13.9	0	0	1	0	13.9
ana	s2	DS	1	24.8	0	0	1	0	24.8
dac1	s2	DS	1	8.6	0	0	1	0	8.6
dac1	s1	G	1	26.7	0	15.8	1	0	42.5
taa5	s1	G	1	433.7	0	11.36	1	0	455.06
lir1	s1	G	3	434.57	0	30.37	3	0	464.94
hak1	s1	G	1	74.2	0	0	1	0	74.2
ana	s1	G	2	49.05	1	11.7	2	0	60.75
mha1	s2	G	2	7.6	1	42.3	2	0	49.9
mrs1	s2	G	1	43.8	0	42.05	1	0	85.85
ala6	s2	G	2	50.65	0	156.03	2	0	206.68
inm1	s1	OD	4	98.9	4	0	4	1	98.9
ala6	s1	OD	2	160.7	1	64.2	2	0	224.9

mha1	s1	OD	14	76.9	14	73.2	14	1	150.1
uss1	s2	OD	19	27.64	18	0	19	0	27.64
taa5	s2	OD	6	88.4	5	0	6	0	88.4
hak1	s2	OD	3	92.3	0	60.9	3	0	153.2
iyj1	s2	OD	4	72.1	4	28.33	4	1	100.43
jib1	s2	OD	4	81.63	4	0	4	1	81.63

2.2 Novice User Group Results for Complex Tasks

The terminology remains the same as Section 2.1 except for c(1/2/3) under TSK which represents complex task (1/2/3).

USR	TSK	PLT	N O/S	T O/S	O/S Fails	T I/S	QRY	TSK Fails	TCT
uss1	c2	DS	1	77.1	0	33	1	0	110.1
jib1	c2	DS	1	60.6	0	21.03	1	0	81.63
hak1	c3	DS	1	21.4	0	63.88	1	0	85.28
hak1	c1	DS	2	54.75	1	24.32	2	0	79.07
taa5	c3	DS	1	194.6	0	38.85	1	0	233.45
mrs1	c3	DS	2	125.4	1	47.4	2	0	172.8
ala6	c2	DS	1	140.5	0	83.34	1	0	233.84
lir1	c3	DS	1	80.1	0	81.78	1	0	161.88
ala6	c1	G	3	60.57	0	72.25	3	0	132.82
inm1	c3	G	1	33.3	1	30.1	1	1	63.4
mrs1	c1	G	1	113.8	0	39.54	1	0	153.34
ana	c3	G	4	36.38	3	37	4	0	73.38
mha1	c1	G	2	41.9	1	66.5	2	0	108.4
iyj1	c3	G	2	26.65	0	127.24	2	0	153.89
dac1	c1	G	5	66.7	2	101.23	5	0	167.93
taa5	c1	G	1	86.66	0	26.6	1	0	113.26
uss1	c3	OD	18	92.6	16	80	18	1	172.6
mha1	c3	OD	25	426.35	24	87.42	25	1	513.77
iyj1	c3	OD	5	33.34	5	85.95	5	1	119.29
inm1	c1	OD	5	67.7	5	0	5	1	67.7
dac1	c2	OD	13	29.17	6	52	13	0	81.17
lir1	c1	OD	15	324.5	15	72.8	15	1	397.3
ana	c2	OD	11	23.08	11	104.18	11	1	127.26
jib1	c1	OD	1	316	1	52.73	1	1	368.73

2.3 Expert User Group Results for Simple Tasks

The terminology remains the same as Section 2.1 for the novice user group.

USR	TSK	PLT	N O/S	T O/S	O/S Fails	T I/S	QRY	TSK Fails	TCT
naj1	s1	DS	1	3.5	0	0	1	0	3.5
kra1	s1	DS	2	5.25	1	0	2	0	5.25
tik1	s2	DS	1	20.7	0	0	1	0	20.7
guan	s1	DS	1	20	0	0	1	0	20
ncu1	s1	DS	1	9.3	0	0	1	0	9.3
atl1	s2	DS	2	12	1	0	2	0	12
faa9	s1	DS	1	8.4	0	0	1	0	8.4
gor1	s2	DS	1	12.7	0	0	1	0	12.7
tik1	s1	G	1	227.6	0	0	1	0	227.6
lay1	s1	G	3	25.7	0	4.6	3	0	30.3
atl1	s1	G	1	108.7	0	288.15	1	0	396.85
wow1	s1	G	2	19.3	0	5.8	2	0	25.1
guan	s2	G	1	24.5	0	0	1	0	24.5
haj1	s1	G	2	35.8	1	5.9	2	0	41.7
faa9	s2	G	2	13.9	0	106.85	2	0	120.75
kra1	s2	G	2	54.75	1	35.76	2	0	90.51
yok2	s1	OD	1	17.1	1	0	1	1	17.1
wow1	s2	OD	4	25.33	3	0	4	0	25.33
yok2	s2	OD	4	52.31	3	0	4	0	52.31
ncu1	s2	OD	5	59.84	5	0	5	1	59.84
haj1	s2	OD	4	103.4	3	105.4	4	0	208.8
gor1	s1	OD	2	49.95	2	0	2	1	49.95
naj1	s2	OD	2	15.85	0	0	2	0	15.85
lay1	s2	OD	3	17.93	2	0	3	0	17.93

2.4 Expert User Group Results for Complex Tasks

The terminology remains the same as Section 2.2 for the novice user group.

USR	TSK	PLT	N O/S	T O/S	O/S Fails	T I/S	QRY	TSK Fails	TCT
tik1	c3	DS	1	42.5	0	69.92	1	0	112.42
lay1	c1	DS	1	126.6	0	45.73	1	0	172.33
kra1	c3	DS	3	60.43	2	45.72	3	0	106.15

ncu1	c2	DS	1	25	0	37.18	1	0	62.18
naj1	c1	DS	1	31.1	0	6.22	1	0	37.32
atl1	c1	DS	1	75.3	0	42.67	1	0	117.97
gor1	c2	DS	1	40	0	31.3	1	0	71.3
faa9	c1	DS	1	20.9	0	35.3	1	0	56.2
lay1	c3	G	1	115.2	0	41.36	1	0	156.56
naj1	c3	G	2	59.9	2	20	2	1	79.9
yok2	c3	G	1	55.5	0	48.32	1	0	103.82
wow1	c3	G	2	24	0	62.96	2	0	86.96
tik1	c1	G	1	54.23	0	41.3	1	0	95.53
guan	c2	G	1	64	0	26.88	1	0	90.88
haj1	c1	G	2	35.25	0	60.52	2	0	95.77
gor1	c1	G	2	39.45	0	29.74	2	0	69.19
yok2	c2	OD	14	48.9	13	48.5	14	1	97.4
guan	c1	OD	8	18.7	3	46.5	8	0	65.2
wow1	c1	OD	2	51.9	1	31.93	2	1	83.83
ncu1	c1	OD	8	19.73	2	60.02	8	1	79.75
kra1	c1	OD	11	85.4	10	42.6	11	1	128
faa9	c3	OD	18	75.3	13	56.71	18	0	132.01
atl1	c3	OD	5	25.66	5	0	5	1	25.66
haj1	c2	OD	3	15.37	3	0	3	1	15.37

2.5 User Feedback Ratings for D-Search

DS represents for D-Search;

VG represents for D-Search in comparison to Google;

VOD represents for D-Search in comparison to the Open Directory.

USR	**GRP**	**Effectiveness**			**Efficiency**			**Successful**			**Helpfulness**		
		DS	**VG**	**VOD**	**DS**	**VG**	**VOD**	**DS**	**VG**	**VOD**	**DS**	**VG**	**VOD**
ala6	N	4	3	5	5	3	5	5	5	5	5	4	5
ana	N	6	4	6	5	4	6	5	4	6	6	5	5
dac	N	5	3	5	4	3	5	4	4	4	5	4	5
hak5	N	6	4	6	5	4	5	4	5	6	5	6	6
inm1	N	4	5	4	3	5	5	3	5	3	2	5	5
iyj1	N	4	3	4	4	4	3	4	5	4	5	4	5
jib	N	4	5	6	6	5	6	5	6	5	6	6	6
lir1	N	5	5	6	6	6	6	5	6	6	6	6	6

mha1	N	6	5	6	6	5	5	5	4	6	5	6	6
mrs1	N	5	5	5	5	2	5	4	5	5	6	6	6
taa5	N	6	6	6	6	6	6	6	6	6	6	6	6
uss1	N	5	4	5	4	3	5	4	3	5	6	4	5
atl1	E	5	4	5	4	6	6	6	4	5	6	5	5
faa9	E	5	5	6	5	4	6	5	4	6	6	6	6
gor1	E	6	4	6	5	4	6	5	5	5	6	6	6
guan	E	5	6	6	6	6	5	5	6	6	5	6	6
haj1	E	3	5	2	4	5	2	4	6	4	4	6	6
kra1	E	4	3	6	5	4	6	4	5	6	5	5	5
lay1	E	5	5	5	3	2	5	5	4	5	6	6	6
naj1	E	5	6	6	6	5	6	5	5	6	6	6	6
ncu1	E	6	6	5	6	3	6	5	4	6	5	4	6
tik1	E	5	4	6	5	2	6	6	6	6	5	6	5
wow1	E	6	5	6	4	4	6	5	3	5	6	5	5
yok2	E	5	4	6	6	6	6	6	6	6	6	6	6

2.6 User Feedback Ratings for D-Persona

DP represents for D-Persona.

USR	GRP	DP Effectiveness	DP Efficiency	DP Successful	DP Helpfulness
ala6	N	2	3	3	2
ana	N	5	6	5	5
dac	N	6	6	6	5
hak5	N	5	5	4	6
inm1	N	4	5	5	5
iyj1	N	6	5	5	5
jib	N	6	6	6	5
lir1	N	6	5	6	6
mha1	N	6	6	6	6
mrs1	N	5	5	4	5
taa5	N	5	4	5	4
uss1	N	4	5	5	5
atl1	E	6	5	5	4
faa9	E	6	6	6	4
gor1	E	5	6	4	5
guan	E	5	1	2	2
haj1	E	5	5	4	4
kra1	E	5	5	6	6

lay1	E	6	6	5	6
naj1	E	6	6	4	6
ncu1	E	6	4	4	6
tik1	E	6	6	5	6
wow1	E	4	5	5	5
yok2	E	6	4	5	6

2.7 Subjective Feedbacks for D-Search & D-Persona

"I think it is a good idea for using directory though I feel it is necessary to get use to it (the way the user interprets with the system) for the service to be of great help. The advantage seems to be the way categories are represented to user and also the personalisation service. The disadvantage or perhaps the difficulty was that some results (categories) didn't seem to completely relevant (though I haven't get into more details)" (atl1).

"The advantages include: using similar concepts from library catalogue search which is consistent, reduces the difficulties in accurately translating user's need into search expressions; personalisation is a very good idea to help user use Web directories as it reflects a very common user demand of using Web directories which is that users do not always need to see the whole classification, instead, they are interested in the classifications they are interested; disadvantages include: it is a new concept which may require some time for users getting familiar with so that it would need more detailed instructions and examples as well as better interface" (jib1).

"Personalisation takes advantages from search engine and the original directory, which is very good. But personalisation could make trouble to people who are not familiar to directory in terms of use. The new search is better than the old one in searching through directory but the interface could still be improved. For example, as simple as Google" (lir1).

"D-Search and D-Persona feature would be very useful. Maybe the descriptions of results can be used to find similar sites in the directory, which could also help the

personalised section too. Disadvantage is if you do not know any site to start off with. Maybe a general [google type] search could aid user in finding first site, maybe from the DMOZ descriptions, as they as much better than google summaries" (kra1).

"I really liked the idea of D-Search and the way of looking in the directory, the idea is perfectly fine but it needs a little bit of improvement. For example, sometimes, the website (used as keyword) typed in the field of search is not recognised by D-Search or not accessible. Also, I think it will be better if there will be a nice design for the directory which can give it the look of a professional work. In general, I really liked the idea and I really hope the portal can be finished soon and we can get access of use it" (haj1).

"D-Search is easy to search for something by just inputting a similar website when you can't think of any keywords. But it might not produce many results when a particular website is entered, so there is no guarantee of a result. It could be improved with better interface and should display better when a matched category is found" (jason).

"The advantages include D-Search and D-Persona are quick and easy in finding results and they are straight to the point when looking for some information such as flights etc. The disadvantages include you have to be accurate with website names and spelling mistakes can delay search" (inm1).

"I think D-Persona is a very good idea of listing the interests and D-Search offers a good way to search topics. In terms of usage D-Search is a bit difficult for the user to get started. If D-Search had also added the prefixes of the website it would be best" (hak5).

"D-Search is a very accurate way of finding something specific and offer relevant objects/items related very closely to your search, since you get the list of all similar websites. D-Persona is a very good idea to personalise your directory. The only thng I think should be looked over, is switching between Google and D-Search in order to make sure the web address is spelt correctly" (mha1).

"The basic idea of D-Search is very useful. But the interface needs improvements. It is not convenient to go to Google first to check the spelling of a website and come back. Users want all-in-one solutions. The D-Persona is a good idea to personalise the directory but sometimes I couldn't find out relevant results to my websites" (mrs1).

“D-Search makes complex searching easy and helps in finding relevant websites easily with more precise results. But it is difficult to remember complete addresses of websites. Http://www or https://www are general words so they should be written as default in the text field. It would be more easy if instead of url, only some keywords about website be written, and system automatically decides about the URL. D-Persona is a very good solution to help me browse the directory as it only shows what I am interested” (uss1).

“D-Search's URL query is very good, D-Persona allows you to customise the directory. However, the categories displayed as results have no descriptions, and the hint phrase of use search engine is too long. In the future, it could have better interface for results, for D-Search, automatically put in http://, more instruction for typing in interested websites” (lay1).

“D-Search & D-Persona are extremely helpful for what they are designed to do. I personally wouldn't need to use them now but I'm sure it will be useful later on in life” (yok2).

“D-Search is a good portal and the idea behind is strong. But I think it has to be integrated with some search engine as well because what if I don't know about a website for a particular topic or thing of interests. D-Persona offers you the opportunity to customise a Web directory which is very good.” (ala6)

“D-Search is easier to get the necessary website without going through selection of results in order to find the sites. It saves times and you don't have to think too hard under which categories something that you are looking for. However, it has a lack of recognition when it comes to finding a particular page of a websites, also it will be good if it will accept the address of a site that normally the browser will accepts, for example, www.caramail.com will automatically be replaced by www.caramail.lycos.fr in the browser but D-Search won't have recognised the first one. D-Persona is good as it helps you browse only the information you are interested not the whole chunk of information” (faa9).

“D-Search and D-Persona are good ideas as relating results (with high-hits) are common user habits of searching. This system enhances this search model. However, it needs sometime to get used to search in this way.” (guan).

“I think D-Search is a very helpful tool to utilise a particular website of a particular

topic when you do not know any websites related to that topic. It is a much faster way to reach the websites relevant to my area of interest. D-Persona is good because it helps you filter the information you are interested. I think that as an improvement, it should allow the user to type in a topic as well in addition to websites. The user could then find its own sub directory that interests him" (taa5).

"Well, the best thing is that you can find websites similar to these that you already know. However, you need a 'start point', you need to use a search engine to find a website's exact path. So it degrades the experience. It would be more that in the case that the address path provided was not found, the system" would suggest you some. Directory, in my opinion is quite useless, but D-Search and D-Persona helps to change my opinion a lot, and they are very simple to find a set of websites with a common topic. (gor1).

"I found D-Search and D-Persona refreshed Web directories. I have come across the word before but never fully knew what in meant. The raw directory structure is hard to navigate, it is easier to use D-Persona to find some. More spacing (1.5 or double) between the results will aid legibility" (ncu1).

"D-Search provides a good amount of high quality websites that have already been checked by volunteers. In this point it is clearly better than Google. On the downside, however, it is slightly more difficult to use and requires more action from the user. Users may not be familiar with the interface, because Google's interface is far simpler. D-Search should ideally implement a more user-friendly interface, with less text to read on the screen." (tik1)

"It was the 1st time I got involved with the use of a directory and now I feel more confident and willing to use it than before because D-Search and D-Persona can offer you faster and more accurate solutions" (dac).

"Good point is D-Search is quick and accurate compared to Google but sometimes you need to make sure the accuracy of the website you are going to use before starting. D-Persona, in the other way, is a very good solution for customise the interface of a Web directory but like D-Search, sometimes a website can not be found in the directory so it needs some recommendations by the system" (ana).

"Good side is, D-Search can help you search related categories quickly and the results

are displayed from A-Z, it is faster than using the original search engine on the directory. However, some websites cannot be found in the directory so they returned no results. Moreover, sometimes it is difficult to know a website before launching a search. My suggestion is to include keyword search at some point. For example, topics. D-Persona is good to use as it simplifies the navigation in Web directory"(zhw3).

"I would say D-Search and D-Persona are both good enhancements for Web directories as the original Web directory is so difficult to use, especially for its search engine. However, I found that sometimes it is not easy to keep your sample website accurate as you have no idea which address is used by the directory, for example, www.mymemory.com or www.mymemory.co.uk. So D-Search should have a mechanism to help you validate your input and make suggestions whenever is possible" (naj1).

Bibliography

A

- About the Virtual Library. In *The WWW Virtual Library*. Retrieved August 19, 2008, from http://vlib.org/admin/AboutVL
- Ackrill, J. L. (1963). *Aristotle's "Categories", and "De Interpretatione"*, Clarendon Aristotle Series. Oxford, England: Oxford Press.
- Adkisson, H. P. (2005). Use of Faceted Classification. *Web Design Practices*. Retrieved August 24, 2008, from http://www.webdesignpractices.com/navigation/facets.html
- Adomavicius, G., & Tuzhilin, A. (2001). Expert-Driven Validation of Rule-BAsed User Models in Personalization Applications. *Data Mining and Knowledge Discovery*, *5*(1-2), 33-58.
- Adomavicius, G., & Tuzhilin, A. (2002). e-Butler: An Architecture of a Customer-Centric Personalization System. *International Journal of Computational Intelligence and Applications*, *2*(3), 313-327.
- Adomavicius, G., & Tuzhilin, A. (2005). Personalization technologies: a process-oriented perspective. *Communications of the ACM*, *48*(10), 83-90.
- Agrawal, R., & Srikant, R. (1995). Mining Sequential Patterns. In *Proceedings of the Eleventh International Conference on Data Engineering*. Taipei, Taiwan.
- Armstrong, R., Freitag, D., & Joachims, T. (1995). WebWatcher: A Learning Apprentice for the World Wide Web. In *Proceedings of the 1995 AAAI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environments*.
- Asnicar, F., & Tasso, C. (1997). ifWeb: a prototype of user model-based intelligent agent for document filtering and navigation in the World Wide Web.

In *Proceedings of the Sixth International Conference on User Modeling*. Chia Laguna, Sardinia.

B

- Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. New York, NY: ACM Press / Addison-Wesley.
- Balabanović, M. (1997). An adaptive Web page recommendation service. In *Proceedings of the first international conference on Autonomous agents* (pp. 378-385). Marina del Rey, CA.
- Balabanović, M., & Shoham, Y. (1997). Fab: content-based, collaborative recommendation. *Communications of the ACM*, *40*(3), 66-72.
- Barrett, R., Maglio, P., & Kellem, D. (1997). WBI: a confederation of agents that personalize the Web. In *Proceedings of the first international conference on Autonomous agents*. Marina del Rey, CA.
- Bates, M. J. (1989). The Design of Browsing and Berrypicking Techniques for the Online Search Interface. *Online Review*, *13*(5), 407-424.
- Battles, M. (2003). *Library: An Unquiet History* (1st ed.). New York, NY: W. W. Norton & Company.
- Beard, D. V., & Walker, J. Q. (1990). Navigational techniques to improve the display of large two-dimensional spaces. *Behavior & Information Technology*, *9*, 451-466.
- Beaudoin, L., Parent, M., & Vroomen, L. C. (1996). Cheops: a compact explorer for complex hierarchies. In *Proceedings of the 7th conference on Visualization '96* (pp. 87-ff). Los Alamitos, CA: IEEE Computer Society Press.
- Bederson, B. B., Shneiderman, B., & Wattenberg, M. (2002). Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies. *ACM Transactions on Graphics*, *21*(4), 833-854.
- Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982). Ask for information

retrieval: Part I. Background and theory. *Journal of Documentation*, *38*(2), 61-71.

- Bell, B. (1992). *Using programming walkthroughs to design a visual language*. Technical Report, Boulder, CO: University of Colorado.
- Bias, R. G. (1991). Walkthroughs: Efficient collaborative testing. *IEEE Software*, *8*(5), 94-95.
- Bias, R. G. (1994). The pluralistic usability walkthrough: Coordinated empathies. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods* (pp. 173-202). New York, NY: John Wiley & Sons.
- Blandford, A. E., & Green, T. R. G. (1997). OSM: an ontology-based approach to usability evaluation. *Workshop on Representations*. Queen Mary College London.
- Blandford, A. E., & Green, T. R. G. (1998). Ontological Sketch Models of Schedule Organisers. In *Proceedings of Workshop on Understanding Work and Designing Artefacts*. University of York.
- Blandford, A. E., & Green, T. R. G. (2001). From tasks to conceptual structures: misfit analysis. In *Proceedings of IHM-HCI2001* (Vol. 2).
- Blandford, A. E., Green, T. R. G., & Connell, I. (2005). Formalising an understanding of user–system misfits. In P. Bastide, P. Palanque, & J. Roth (Eds.), *Proceedings of EHCI-DSVIS 2004* (pp. 253-270). Springer.
- Blandford, A. E., & Young, R. M. (1996). Specifying user knowledge for the design of interactive systems. *Software Engineering Journal*, *11*(6), 323-333.
- Bliss, H. E. (1940). *A bibliographic classification, extended by systematic auxiliary schedules for composite specification* (Vols. 1-4). New York, NY: H.W. Wilson.
- Bliss, H. E. (1910). A modern classification for libraries, with simple notation, mnemonics, and alternatives. *Library Journal*, *35*, 351-358.
- Bliss, H. E. (1935). *A system of bibliographic classification*. New York, NY: H.W. Wilson.

- Bonett, M. (2001). Personalization of Web Services: Opportunities and Challenges. *Ariadne*, Issue 28. Retrieved August 24, 2008, from http://www.ariadne.ac.uk/issue28/personalization/intro.html
- Borgman, C. L. (1986). Why are online catalogs hard to use? Lessons learned from information-retrieval studies. *Journal of the American Society for Information Science*, *37*(6), 387-400.
- Borgman, C. L. (1996). Why are online catalogs still hard to use? *Journal of the American Society for Information Science*, *47*(7), 493-503.
- Bosshard, M. (2001, August 17). Services tailored to suit your personal needs. *Credit Suisse*. Retrieved September 25, 2008, from http://emagazine.credit-suisse.com/app/article/index.cfm?fuseaction=OpenArticle&aoid=3339&lang=en
- Boswell, W. (2004). Web Directory – What Is a Web Directory. *About.com*. Retrieved August 19, 2008, from http://websearch.about.com/od/enginesanddirectories/a/subdirectory.htm
- Breese, J., Heckerman, D., & Kadie, C. (1998). Empirical Analysis of Predictive Algorithms for Collaborative Filtering. In *Proceedings of Fourteenth Conference on Uncertainty in Artificial Intelligence*. Madison, WI.
- Brewington, B. E., & Cybenko, G. (2000). How dynamic is the Web? In *Proceedings of the 9th international World Wide Web conference on Computer networks: the international journal of computer and telecommunications networking* (pp. 257-276). Amsterdam, The Netherlands.
- Brittain, J. (1971). *Information and its Users - A Review with Special Reference to Social Sciences*. New York, NY: Wiley Interscience.
- Brittain, J. (1975). Information needs and applications or results of user studies. In A. Debons & W. J. Cameron (Eds.), *Perspectives in Information Science* (pp. 425-447). Leyden, NL: Noordhoff International Publishers.
- Broughton, V. (2002). Faceted classification as a basis for knowledge organization in a digital environment: the bliss bibliographic classification as a model for vocabulary management and the creation of multidimensional knowledge structures. *The New Review of Hypermedia and Multimedia*, *7*(1),

67-102.

- Buchner, A. G., & Mulvenna, M. D. (1998). Discovering internet marketing intelligence through online analytical web usage mining. *SIGMOD Record*, *27*(4), 54-61.
- Buckley, C., & Salton, G. (1995). Optimization of relevance feedback weights. In *Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval*. Seattle, Washington.
- Byström, K., & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing and Management*, *31*(2), 191-213.

C

- Carriere, J., & Kazman, R. (1995). Research report: Interacting with huge hierarchies: beyond cone trees. In *Proceedings of the 1995 IEEE Symposium on Information Visualization* (p. 74). Washington, DC: IEEE Computer Society.
- Chan, L. M. (1986). *Library of Congress Subject Headings: Principles and Application*. Littleton, CO: Libraries Unlimited.
- Chan, L. M. (1995). Classification, Present and Future. *Cataloging & Classification Quarterly*, *21*(2), 5-17.
- Chan, L. M. (1999). *A Guide to the Library of Congress Classification*. Englewood, CO: Libraries Unlimited.
- Chan, L. M. (2000). Exploiting LCSH, LCC, and DDC To Retrieve Networked Resources Issues and Challenges. In *Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web*.
- Chan, L. M. (2003). *Library of Congress Classification in a New Setting: Beyond Shelfmarks*. The Library of Congress. Retrieved August 24, 2008, from http://www.loc.gov/cds/chanarticle.html
- Chen, L., & Sycara, K. (1998). WebMate : A Personal Agent for Browsing and

Searching. In *Proceedings of the second international conference on Autonomous agents*. Minneapolis, Minnesota.

- Chellappa, R. K., & Sin, R. G. (2005). Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma. *Information Technology and Management* , *6*(2-3), 181-202.
- Chowdhury, S., & Chowdhury, G. G. (2004). Using DDC to create a visual knowledge map as an aid to online information retrieval. *Paper presented at the ISKO8 conference*. London.
- Clementi, L. (2007). *Information Visualization Focus and Context*. Switzerland: University of Fribourg.
- Coates, E. J. (1960). *The British Catalogue of Music Classification / compiled for the Council of the British National Bibliography Ltd*. London: Council of the British National Bibliography.
- Cochrane, P., & Johnson, E. (1996). Visual Dewey: DDC in a Hypertextual Browser for the Library User. *In Knowledge Organization and Change: Proceedings of the 4th International ISKO Conference*. In R. Green (Ed.), (pp. 95-106). Frankfurt: Indeks Verlag.
- Collaborative filtering. (2008, September 23). *Wikipedia, The Free Encyclopedia*. Retrieved October 1, 2008, from http://en.wikipedia.org/w/index.php?title=Collaborative_filtering&oldid=240385501
- Connell, I., Blandford, A. E., & Green, T. R. G. (2002). Usability of a Music Digital Library: an OSM Case Study. *CASSM Project*. Retrieved August 24, 2008, from http://web4.cs.ucl.ac.uk/uclic/annb/docs/ISMIR-OSM.pdf
- Connell, I., Green, T. R. G., & Blandford, A. E. (2003). Ontological Sketch Models: Highlighting User-System Misfits. In E. O'Neill, P. Palanque, & P. Johnson (Eds.), *People and Computers XVII, Proceedings of HCI'03* (pp. 163-178). Springer.
- Cooley, R., Mobasher, B., & Srivastava, J. (1999). Data preparation for mining world wide web browsing patterns. *Information Systems*, *1*(1), 5-32.

- Cortes, C., Fisher, K., Pregibon, D., & Rogers, A. (2000). Hancock: a language for extracting signatures from data streams. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 9-17). Boston, MA.
- Culnan, M. J. (1983). Environmental scanning : The effects of task complexity and source accessibility on information gathering behavior. *Decision Sciences*, *14*(2), 194-206.
- Cutter, C. A. (1962). *Rules for a Dictionary Catalog* (W. P. Cutter, Ed.) (4th ed.). Washington, D.C.: Government Printing Office.

D

- Denton, W. (2003, October). Putting Facets on the Web: An Annotated Bibliography. *Miskatonic University Press*. Retrieved August 24, 2008, from http://www.miskatonic.org/library/facet-biblio.html
- Denton, W. (2003, November). How to Make a Faceted Classification and Put It On the Web. *Miskatonic University Press*. Retrieved August 24, 2008, from http://www.miskatonic.org/library/facet-web-howto.html
- Dervin, B., & Nilan, M. S. (1986). INFORMATION NEEDS AND USES. In *Annual Review of Information Science and Technology (ARIST)* (Vol. 21, pp. 3-33). White Plains, NY: Knowledge Industry Publications.
- Dewey Decimal Classification. (2008, September 24). *Wikipedia, The Free Encyclopedia*. Retrieved 15:50, September 25, 2008, from http://en.wikipedia.org/w/index.php?title=Dewey_Decimal_Classification&oldid=240715964
- Dewey, M. (1979). *Decimal Classification and Relative Indes* (19th ed.). Albany, New York, NY: Forest press.
- Dewey Services. *OCLC - Cataloging and Metadata*. Retrieved August 24, 2008, from http://www.oclc.org/dewey/

- Donelson, W. C. (1978). Spatial management of information. *ACM SIGGRAPH Computer Graphics*, *12*(3), 203-209.
- Dot-com bubble. (2008, August 18). *Wikipedia, The Free Encyclopedia*. Retrieved August 21, 2008, from http://en.wikipedia.org/w/index.php?title=Dot-com_bubble&oldid=232686980
- Drabenstott, K. M. (1989). Subject Searching Strategies for Online Catalogues Through the Dewey Decimal Classification. In C. R. Hildreth (Ed.), *Online Library Catalogues: Essentail Questions and Future Prospects* (pp. 61-83). London: Library Association Publishing.
- Drengson, A. (1996). Editorial: How Rigid the Hierarchy? *Trumpeter*, *13*(2).
- Dumas, J. S., & Redish, J. C. (1993). *A practical guide to usability testing*. Norwood, NJ: Ablex Publishing.
- Duncan, E. (157). A faceted approach to hypertext. In R. McAleese (Ed.), *Hypertext: theory into practice*. London: Intellect.
- Dykstra, M. E. (2004). Information Organization. *Encyclopedia Americana Online*.

E

- Ellis, D. (1989). A Behavioural Approach to Information Retrieval System Design. *Journal of Documentation*, *45*(3), 171-212.
- Ellis, D., Cox, D., & Hall, K. (1993). A Comparison of the Information-seeking Patterns of Researchers in the Physical and Social Sciences. *Journal of Documentation*, *49*(4), 356-369.
- Ellis, D., & Haugan, M. (1997). Modelling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of Documentation*, *53*(4), 384-403.
- Ellis, D., & Vasconcelos, A. C. (1999). Ranganathan and the Net: Using facet analysis to search and organize the World Wide Web. *Aslib Proceedings*, *51*(1),

3-10.

- Ellis, D., & Vasconcelos, A. C. (2000). The relevance of Facet Analysis for World Wide Web subject organisation and searching. *Journal of Internet Cataloging*, *2*(3/4), 97-114.

F

- Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, *35*(3), 379-383.
- Foltz, P. W., & Dumais, S. T. (1992). Personalized information delivery: An analysis of information filtering methods. *Communications of the ACM*, *35*(12), 51-60.
- Foskett, D. J. (1972). Facet analysis. In A. Kent & H. Lancour (Eds.), *Encyclopedia of library and information science* (pp. 338-346). New York, NY: Marcel Dekker.
- Foskett, A. C. (2000). The future of faceted classification. In R. Marcella & A. Maltby (Eds.), *The Future of Classification* (pp. 69-80). Aldershot: Gower.
- Freedman, D., & Weinberg, G. M. (1990). *Handbook of Walkthroughs, Inspections, and Technical Reviews: Evaluating Programs, Projects, and Products*. New York, NY: Dorset House.

G

- Goecks, J., & Shavlik, J. (1999). Automatically Labeling Web Pages Based on Normal User Actions. In *Proceedings of the IJCAI-99 Workshop on Machine Learning for Information Filtering*.
- Gilb, T., & Graham, D. (1993). *Software Inspection* (S. Finzi, Ed.). Boston, MA: Addison-Wesley.

- Google Web Directory Help. *Google*. Retrieved August 24, 2008, from http://www.google.com/dirhelp.html
- Google Milestones. *Corporte Information*. Retrieved September 25, 2008, from http://www.google.com/corporate/history.html
- Green, T. R. G., & Benyon, D. (1995). Displays as data structures: entity-relationship models of information artefacts. *INTERACT 1995* (pp. 197-202).
- Green, T. R. G., & Benyon, D. (1996). The skull beneath the skin: entity-relationship models of information artifacts. *International Journal of Human-Computer Studies*, *44*(6), 801-828.
- Gunn, C. (1992). Visualizing hyperbolic space. *Computer Graphics and Mathematics*, 299-313.

H

- Han, E., Boley, D., Gini, M., Gross, R., Hastings, K., Karypis, G., et al. (1998). WebACE: a Web agent for document categorization and exploration. In *Proceedings of the second international conference on Autonomous agents*. Minneapolis, Minnesota.
- Han, J., & Fu, Y. (1995). Discovery of multiple-level association rules from large databases. In *Proceedings of the 21th International Conference on Very Large Data Bases*.
- Hand, D., Mannila, H., & Smyth, P. (2001). *Principles of Data Mining*. Cambridge, MA: The MIT Press.
- Harman, D. (1995). Overview of the Second Text Retrieval Conference (TREC-2). *Information Processing and Management*, *31*(3), 271-289.
- Hart, P. J., & Rice, R. E. (1991). Using information from external databases: contextual relationships of use, access method, task, database type, organizational differences, and outcomes. *Information Processing and Management*, *27*(5), 461-479.

- Hawley, D. (2005, January 12). Rise and Fall of Generic Directories. *Search Engine Watch Forums*. Retrieved August 24, 2008, from http://forums.searchenginewatch.com/showthread.php?t=3736
- He, P., Koh, S., & Shumin, W. (2004). Faceted Classification. *LIBR 517: Document and Subject Analysis*. Retrieved April 12, 2003, from http://www.slais.ubc.ca/COURSES/libr517/02-03-wt2/projects/faceted/index.htm
- He, G. (1990). *Theory of Library Information Classification (Trans)*. Taipei: Sanmin Press.
- Hom, J. (1998, June 10). Guideline Checklists . *The Usability Methods Toolbox*. Retrieved August 24, 2008, from http://jthom.best.vwh.net/usability/guidline.htm
- Herot, C. F., Carling, R., Friedell, M., & Kramlich, D. (1980). A Prototype Spatial Data Management System. *ACM SIGGRAPH Computer Graphics*, *14*(3), 63-70.
- Hjørland, B. Lifeboat for Knowledge Organization. *Lifeboat for Knowledge Organization*. Retrieved August 24, 2008, from http://www.db.dk/bh/Lifeboat_KO/home.htm
- Hiom, D. (1998). Mapping classification notes: document notes.
- How does the Yahoo! Directory differ from Yahoo! Search? *Yahoo! Directory.* Retrieved August 19, 2008, from http://help.yahoo.com/l/us/yahoo/directory/basics/basics-03.html
- How to suggest a site to the Open Directory. *Open Directory Project*. Retrieved August 24, 2008, from http://www.dmoz.org/add.html
- Hudon, M. (2000). Innovation and tradition in knowledge organisation schemes on the Internet, or, finding one’s way in the virtual library. In C. Beghtol, L. C. Howarth, & N. J. Williamson (Eds.), *Dynamism and Stability in Knowledge Organisation: Proceedings of the 6th International ISKO Conference* (pp. 35-40). Ergon Verlag, Toronto.

I

- Information technology - Specification and standardization of data elements - Part 2: Classification for data elements (1st ed.). (2000, February 15). *International Standard*. Retrieved August 19, 2008 from http://metadata-standards.org/11179-2/ISO-IEC_11179-2_2000_IS_E.pdf
- Ingwersen, P. (1992). *Information Retrieval Interaction*. London: Taylor Graham.
- Intner, S. S., & Weihs, J. (1996). *Standard Cataloging for School and Public Libraries* (2nd ed.). Englewood, CO: Libraries Unlimited.
- Introduction to Dewey Decimal Classification. (2003). Retrieved August 24, 2008, from http://www.oclc.org/dewey/versions/ddc22print/intro.pdf
- *ISO/IEC 11179-2:2005*. (2005). ISO. Retrieved August 24, 2008, from http://standards.iso.org/ittf/PubliclyAvailableStandards/c035345_ISO_IEC_11179-2_2005(E).zip
- Irwin, S. (2001, March 6). *Classification Theory and the Internet: A move toward Mulitdimensional Classification*. University of Denver. Retrieved August 24, 2008, from http://www.du.edu/lisa/students/student_papers/Irwin.htm
- ISO 9241-11:1998. (1998). ISO. Retrieved August 24, 2008, from http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=16883
- Iyer, H. (1995). *Classificatory Structures: Concepts, relations and representation*. Frankfurt: Indeks Verlag.
- Iyer, H., & Giguere, M. (1995). Towards designing an expert system to map mathematics classificatory structures. *Knowledge Organization*, *22*(3-4), 141-147.

J

- Jeffries, R., Miller, J. R., Wharton, C., & Uyeda, K. (1991). User interface evaluation in the real world: a comparison of four techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology* (pp. 119-124). New York, NY: ACM Press.
- Jenkins, C., Jackson, M., Burden, P., & Wallis, J. (1998). Automatic classification of Web resources using Java and Dewey decimal classification. *Computer Networks and ISDN Systems*, *30*(7), 646-648.
- Jeong, C., & Pang, A. (1998). Reconfigurable Disc Trees for Visualizing Large Hierarchical Information Space. In *Proceedings of the 1998 IEEE Symposium on Information Visualization* (pp. 19-25). Washington, DC: IEEE Computer Society.
- Joachims, T., Freitag, D., & Mitchell, T. (1997). WebWatcher: A Tour Guide for the World Wide Web. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*. Nagoya, Japan.
- Johnson, B., & Shneiderman, B. (1991). Tree-Maps: a space-filling approach to the visualization of hierarchical information structures. In *Proceedings of the 2nd conference on Visualization '91* (pp. 284-291). San Diego: IEEE Computer Society Press.
- Johnson, B. (1992). TreeViz: treemap visualization of hierarchically structured information. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 369-370). New York, NY: ACM Press.
- Jones, K. P. (1990). Informatics 10: Prospects for Intelligent Retrieval. In *Proceedings of a conference jointly sponsored by Aslib, the Aslib Informatics Group and the Information Retrieval Specialist Group of the British Computer Society, Kings College, Cambridge*. London: Aslib.
- Jungmeister, W., & Turo, D. (1992). *Adapting Treemaps to Stock Portfolio Visualization (1992)*. Technical Report, University of Maryland.

K

- Kerber, R. (1999, April 19). Lycos To Turn Ranking System Over To Users Joining "Open Directory Project" Run By Netscape. *The Boston Globe*.
- Kahn, M. J., & Prail, A. (1994). Formal usability inspections. In *Usability Inspection Methods* (pp. 203-232). New York, NY: John Wiley & Sons.
- Kane, M. (1999). Yahoo! asks sites to pay for privilege. *ZDNet News*. Retrieved August 19, 2008, from http://news.zdnet.com/2100-9595_22-513694.html
- Keith, A. (2002). Information Visualisation.
- Kim, K., Sin, S. J., Park, S., Zhu, X., & Polparsi, J. (2006). Facet Analyses of Categories Used in Web Direcrtories: A Comparative Study. In *IFLA (International Federation of Library Associations) 2006 Conference*. Seoul, Korea.
- Kleiberg, E., van de Wetering, H., & Van Wijk, J. J. (2001). Botanical Visualization of Huge Hierarchies. In *Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01)* (p. 87). Washington, DC: IEEE Computer Society.
- Koch, T. (1997). *The role of classification schemes in Internet resource description and discovery*. DESIRE: Project Deliverable.
- Koch, T., & Day, M. (1997). *The role of classification schemes in Internet resource description and discovery*. Retrieved August 24, 2008, from http://www.ub2.lu.se/desire/radar/reports/D3.2.3/
- Koch, T., & Vizine-Goetz, D. (1998). *Automatic Classification and Content Navigation Support for Web Services - DESIRE II Cooperates with OCLC*. Annual Review of OCLC Research 1998, OCLC.
- Koch, T. (2000). Quality controlled subject gateways: definitions, typologies, empirical overview. *Online Information Review*, *24*(1), 24-34.
- Koike, H., & Yoshihara, H. (1993). Fractal Approaches for Visualizing Huge Hierarchies. In *Proceedings of the 1993 IEEE Symposium on Visual Languages* (pp. 55-60).

- Kuhlthau, C. C. (1991). Inside the search process: information seeking from the user's perspective. *Journal of the American Society of Information Science*, *42*(5), 361-371.
- Kuhlthau, C. C. (1993). A principle of uncertainty for information seeking. *Journal of Documentation*, *49*(4), 339-355.
- Kuhlthau, C. C. (1994). *Seeking Meaning: A Process Approach to Library and Information Services*. Norwood, NJ: Ablex Publishing.
- Kuhlthau, C. C. (1997). The influence of uncertainty on the information seeking behavior of a securities analyst. In *Proceedings of an international conference on Information seeking in context* (pp. 268-274). London: Taylor Graham.
- Kwasnik, B. H. (1992a). The role of classification in knowledge representation and discovery. *Library Trends*, *48*(1), 22-47.
- Kwasnik, B. H. (1992b). The role of classification structures in reflecting and building theory. In R. Fidel, B. H. Kwasnik, & P. Smith (Eds.), *Advances in classification research* (Vol. 3, pp. 63-81). Medford, NJ: Learned Information, for the American Society for Information Science.
- Kwasnik, B. H., & Rubin, V. L. (2003). Streching conceptual structures in classifications across languages and cultures. *Cataloging & Classification Quarterly*, *37*(1/2), 33-47.

L

- Lamping, J., Rao, R., & Pirolli, P. (1995). A Focus+Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 401-408). ACM Press.
- Lee, H. L., & Olson, H. A. (2005). Hierarchical navigation: An exploration of Yahoo! directories. *Knowledge Organization*, *32*(1), 10-24.
- Leung, Y. K. (1989). Human-computer interface techniques for map based

diagrams. In *Proceedings of the third international conference on human-computer interaction on Designing and using human-computer interfaces and knowledge based systems* (pp. 361-368). Boston.

- Leung, Y. K., & Apperley, M. D. (1994). A review and taxonomy of distortion-oriented presentation techniques. *ACM Transactions on Computer-Human Interaction*, *1*(2), 126-160.
- Lewis, J. R. (1994). Sample sizes for usability studies: Additional considerations. *Human factors*, *36*(2), 368-378.
- Lewis, J. R. (2006). Sample sizes for usability tests: mostly math, not magic. *interactions*, *13*(6), 29-33.
- Librarians of Congress. (2006, March 30). *Jefferson's Legacy: A Brief History of the Library of Congress*. Retrieved August 24, 2008, from http://www.loc.gov/loc/legacy/librs.html
- Library classification. (2008, September 18). *Wikipedia, The Free Encyclopedia*. Retrieved September 25, 2008, from http://en.wikipedia.org/w/index.php?title=Library_classification&oldid=239188401
- Library of Congress Classification. (2008, August 28). *Wikipedia, The Free Encyclopedia*. Retrieved September 25, 2008, from http://en.wikipedia.org/w/index.php?title=Library_of_Congress_Classification&oldid=234758224
- Library of Congress Classification Outline. *Cataloging and Acquisitions (Library of Congress)*. Retrieved August 25, 2008, from http://www.loc.gov/catdir/cpso/lcco/
- Lieberman, H. (1995). Letizia: An Agent that assists web browsing. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95)*.
- Lieberman, H. (1997). Autonomous interface agents. In *Proceedings of ACM Conference on Computers and Human Interaction (CHI'97)*. Atlanta, USA.
- Lieberman, H., Van Dyke, N. W., & Vivacqua, A. (1999). Let's browse: a

collaborative Web browsing agent. In *Proceedings of the 4th international conference on Intelligent user interfaces (IUI'99)*. New York, NY.

- Lin, W. C. (2006). The Concept and Applications of Faceted Classification. *Journal of Educational Media & Library Sciences*, *44*(2), 153-171.
- Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969-1980. *Psychological Bulletin*, *90*, 125-152.

M

- Mack, R. L., & Montaniz, F. Observing, predicting, and analyzing usability problems. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods* (pp. 293-336). New York, NY: John Wiley & Sons.
- Malt, B. C. (1995). Category coherence in cross-cultural-perspective. *Cognitive Psychology*, *29*(2), 85-148.
- Maltby, A., & Gill, L. (1979). *The Case for Bliss*. London: Clive Bingley.
- Manber, U., Patel, A., & Robison, J. (2000). The business of personalization: experience with personalization of Yahoo! *Communications of the ACM*, *43*(8), 35-39.
- Mannila, H., Toivonen, H., & Verkamo, I. A. (1995). Discovering frequent episode in sequences. In *Proceedings of the First International Conference on Knowledge Discovery and Data Mining (KDD-95)*.
- Mannila, H., Toivonen, H., & Verkamo, I. A. (1997). Discovery of Frequent Episodes in Event Sequences. *Data Mining and Knowledge Discovery*, *1*(3), 259-289.
- Maple, A. (1995). Faceted Access: A Review of The Literature. *Presented during the Music Library Association Annual Meeting*. Retrieved August 24, 2008, from http://library.music.indiana.edu/tech_s/mla/facacc.rev
- March, J., & Simon, H. (1967). *Organizations* (2nd ed.). New York, NY: Wiley.
- Mason, M. K. How Useful are the Library of Congress Subject Headings.

Retrieved August 24, 2008, from http://www.moyak.com/researcher/resume/papers/lcsh.html

- Matveyeva, S. J. (2002). A Role for Classification: The Organization of Resources on the Internet. *MLA Forum*, *1*(2). Retrieved from http://www.mlaforum.org/volumeI/issue2/roleClassification.html

- McGowan, J. P., Kushmerick, N., & Smyth, B. (2002). Who Do You Want to Be Today? Web Personae for Personalised Information Access. In *Proceedings of the Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*.

- Miksa, F. L. (1998). *The DDC, the Universe of Knowledge, and the Post-Modern Library* (2nd ed.). Albany, New York, NY: Forest press.

- Mitchell, J. S., & Vizine-Goetz, D. (2001). A Research Agenda for Classification. *Dewey Decimal Classification Research*. Retrieved August 24, 2008, from http://worldcat.org/arcviewer/1/OCC/2003/07/31/0000003953/viewer/file1.html

- Mladeni, D. (1998). *Personal WebWatcher: design and implementation*. Technical Report, Ljubljana, Slovenia: Department for Intelligent Systems, J. Stefan Institute.

- Mobasher, B., Cooley, R., & Srivastava, J. (2000a). Automatic Personalization Through Web Usage Mining. *Communications of the ACM*, *43*(8), 142-151.

- Mobasher, B., Dai, H., Luo, T., Sung, Y., Nakagawa, M., & Wiltshire, J. (2000b). Discovery of Aggregate Usage Profiles for Web Personalization. In *Proceedings of the Web Mining for E-Commerce Workshop (WebKDD'2000), held in conjunction with the ACM-SIGKDD Conference on Knowledge Discovery in Databases (KDD'2000)*. Boston, MA.

- Mohammadi-Aragh, M. J., & Jankun-Kelly, T. J. (2005). MoireTrees: Visualization and Interaction for Multi-Hierarchical Data. *EuroVis 2005* (pp. 231-238).

- Monk, A., Walsh, P., & Dix, A. (1988). A comparison of hypertext, scrolling and folding as mechanisms for program browsing. In D. M. Jones & R. Winder

(Eds.), *People and computers IV* (pp. 421-435). Cambridge, England: Cambridge University Press.

- Morita, M., & Shinoda, Y. (1994). Information Filtering Based on User Behaviour Analysis and Best Match Text Retrieval. *In Proceedings of the 17th Annual International ACM-SIGIR Conference* (pp. 272-281). Dublin, Ireland: ACM.Springer.
- Mundie, D. (1995). CyberDewey: A Catalogue for the World Wide Web. *CyberDewey*. Retrieved August 24, 2008, from http://library.tedankara.k12.tr/dewey/index.html
- Munzner, T., & Burchard, P. (1995). Visualizing the Structure of the World Wide Web in 3D Hyperbolic Space. In *Proceedings of VRML '95* (pp. 33-38). San Diego.
- Munzner, T., Guimbretiere, F., Tasiran, S., Zhang, L., & Zhou, Y. (2003). TreeJuxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility. *ACM Transactions on Graphics*, *22*(3), 453-462.

N

- Nielsen, J. (1989). Usability engineering at a discount. In G. Salvendy & M. J. Smith (Eds.), *Designing and Using Human-Computer Interfaces and Knowledge Based Systems* (pp. 394-401). Amsterdam, the Netherlands: Elsevier Science Publishers.
- Nielsen, J. (1990). Paper versus computer implementations as mockup scenarios for heuristic evaluation. In *Proceedings of IFIP INTERACT'90 Third International Conference Human-Computer Interaction (Cambridge, U.K., August 27-31)* (pp. 315-320).
- Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of ACM CHI'90 Conference* (pp. 249-256). Seattle, WA, April 1- 5.
- Nielsen, J. (1992). The Usability Engineering Life Cycle. *Computer*, *25*(3), 12-22.

- Nielsen, J. (1993). *Usability Engineering*. Boston, MA: Academic Press.
- Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *Proceedings of ACM INTERCHI'93 Conference* (pp. 206-213). Amsterdam, The Netherlands.
- Nielsen, J., & Philips, V. L. (1993). Estimating the relative usability of two interfaces: Heuristic, formal, and empirical methods compared. In *Proceedings ACM/IFIP INTERCHI'93 Conference (Amsterdam, The Netherlands, April 24-29)* (pp. 214-221).
- Nielsen, J. (1994). Heuristic evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods* (pp. 25-64). New York, NY: John Wiley & Sons.
- Nielsen, J., & Mack, R. L. (Eds.). (1994). *Usability Inspection Methods*. New York, NY: John Wiley & Sons.
- Nielsen, J. (1995). Usability Inspection Methods. In *CHI 95 Conference Companion* (pp. 377-378).
- Nielsen, J. (1998, October 4). Personalization is Over-Rated. *Jakob Nielsen's Alertbox*. Retrieved August 24, 2008, from http://www.useit.com/alertbox/981004.html
- Nielsen, J. (2000, March 19). Why You Only Need to Test With 5 Users. *Jakob Nielsen's Alertbox*. Retrieved August 24, 2008, from http://www.useit.com/alertbox/20000319.html
- Nielsen, J. (2001, February 18). Success Rate: The Simplest Usability Metric. *Jakob Nielsen's Alertbox*. Retrieved August 24, 2008, from http://www.useit.com/alertbox/20010218.html
- Nielsen, J. Summary of Usability Inspection Methods. *useit.com*. Retrieved August 24, 2008, from http://www.useit.com/papers/heuristic/inspection_summary.html
- Nielsen, J. (2003, January 20). Recruiting Test Participants for Usability Studies. *Jakob Nielsen's Alertbox*. Retrieved August 24, 2008, from http://www.useit.com/alertbox/20030120.html

- Nielsen, J. (2006, June 26). Quantitative Studies: How Many Users to Test? *Jakob Nielsen's Alertbox*. Retrieved October 3, 2008, from http://www.useit.com/alertbox/quantitative_testing.html
- Norman, D. A. (1988). Psychology of Everyday Action. In *The Design of Everyday Things* (pp. 45-46). New York, NY: Basic Book.

O

- Olson, H. A., & Ward, D. B. (1998). Charting a journey across knowledge domains: feminism in the Dewey decimal classification. In W. M. el-Hadi, J. Maniez, & A. S. Pollitt (Eds.), *Structure and Relations in Knowledge Organization: Proceedings of the 5th International ISKO Conference* (pp. 238-244). Ergon Verlag, Würzburg.
- Olson, H. A. (2001). Sameness and difference: A cultural foundation of classification. *Library Resources & Technical Services*, *45*(3), 115-122.
- Ong, T., Legett, J. J., & Yun, U. (2005). Proceedings of the Sixth Mexican International Conference on Computer Science. In *Proceedings of the Sixth Mexican International Conference on Computer Science* (pp. 31-40). Washington, DC: IEEE Computer Society.
- Oostendorp, K. A., Punch, W. F., & Wiggins, R. W. (1994). A Tool for Individualizing the Web. In *Proceedings of the Second International Conference on the World Wide Web*. Chicago, IL.
- Open Directory Project. (2008, August 21). *Wikipedia, The Free Encyclopedia*. Retrieved August 21, 2008, from http://en.wikipedia.org/w/index.php?title=Open_Directory_Project&oldid=233223795
- ONLINE-ADS>> Search engines are dying. (1997, November 10). *Online Advertising Discussion List*. Retrieved August 24, 2008, from http://www.o-a.com/archive/1997/November/0024.html
- Oostendorp, K. A., Punch, W. F., & Wiggins, R. W. (1994). A Tool for Individualizing the Web. In *Proceedings of the Second International Conference*

on the World Wide Web (pp. 49-59). Chicago, Illinois.

P

- Palay, A. J., & Fox, M. S. (1981). Browsing through Databases. In R. N. Oddy (Ed.), *Information Retrieval Research* (pp. 310-324). London: Butterworth-Heinemann.
- Payne, C. (2000, November 22). Everything You Need to Know About Personalization. *Web Developer's Virtual Library*. Retrieved August 24, 2008, from http://www.wdvl.com/Authoring/ASP/Personalization/index.html
- Pazzani, M. J., Muramatsu, J., & Billsus, D. (1996). Syskill & Webert: Identifying interesting web sites. In *In Proceedings of the Thirteenth National Conference on Artificial Intelligence*. Palo Alto, CA.
- Pazzani, M. J. (1999). A Framework for Collaborative, Content-Based and Demographic Filtering. *Artificial Intelligence Review*, *13*(5-6), 393-408.
- Peppers, D., & Rogers, M. (2001, December). Are Personalization and Privacy Really at Odds?, Customer Service Article - Inc. Article. *INC.com*. Retrieved October 1, 2008, from http://www.inc.com/articles/2001/12/23773.html
- Perfetti, C., & Landesman, L. (2001, June 18). Eight is Not Enough. *User Interface Engineering*. Retrieved October 3, 2008, from http://www.uie.com/articles/eight_is_not_enough/
- Phillips, M., & Gunn, C. (1992). Visualizing hyperbolic space: unusual uses of 4x4 matrices. In *Proceedings of the 1992 symposium on Interactive 3D graphics* (pp. 209-214). New York, NY: ACM Press.
- Pollitt, A. S. (1998). The application of Dewey Classification in a view-based searching OPAC. In *Proceedings of the Fifth international conference of the International Society for Knowledge Organization*. Würzburg: Ergon Verlag.
- Polson, P., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: A method for theory-based evaluation of user interfaces.

International Journal of Man-Machine Studies, *36*(5), 741-773.

- Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, *14*(3), 130-137.
- Pretschner, A., & Gauch, S. (1999). *Personalization on the Web*. Technical Report, Information and Telecommunication Technology Center: Department of Electrical Engineering and Computer Science, The University of Kansas.

Q

- Qin, J., & Paling, S. (2001). Converting a controlled vocabulary into an ontology: the case of GEM. *Information Research*, *6*(2).

R

- Ranganathan, S. R. (1951). *Philosophy of Library Classification*. Copenhagen: Munksgaard.
- Ranganathan, S. R. (1960). *Elements of Library Classification* (2nd ed.). New York, NY: Asia Publishing House.
- Ranganathan, S. R. (1967). *Prolegomena to Library Classification* (3rd ed.). New York, NY: Asia Publishing House.
- Resnick, P., Lacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. Chapel Hill, North Carolina.
- RFC2235. (1997, November). In *The Internet Engineering Task Force*. Retrieved August 19, 2008, from http://www.ietf.org/rfc/rfc2235.txt
- Rich, E. (1998). User modeling via stereotypes. In *Readings in intelligent user interfaces* (pp. 329-342). San Francisco, CA: Morgan Kaufmann Publishers.
- Riecken, D. (2000). Introduction: personalized views of personalization. *Communications of the ACM*, *43*(8), 26-28.

- Robertson, G., Mackinlay, J. D., & Card, S. K. (1991). Cone trees: Animated 3D visualizations of hierarchical information. In *Proceedings of ACM SIGCHI conference on Human Factors in Computing Systems '91* (pp. 189-194). New York, NY: ACM Press.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Applied Linguistics*, *22*(1), 27-57.
- Rowley, D. E., & Rhoades, D. G. (1992). The cognitive jogthrough: a fast-paced user interface evaluation procedure. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 389-395).
- Rucker, J., & Polanco, M. (1997). Siteseer: personalized navigation for the Web. *Communications of the ACM*, *40*(3), 73-76.

S

- Saeed, H., & Chaudhry, A. S. (2001). Potential of bibliographic tools to organise knowledge on the Internet: a review of the use of Dewey decimal classification scheme for organising Web-based information resources. *Knowledge Organization*, *28*(1).
- Saeed, H., & Chaudhry, A. S. (2002). Using Dewey decimal classification scheme (DDC) for building taxonomies for knowledge organisation. *Journal of Documentation*, *58*(5), 575-583.
- Salton, G., & McGill, M. J. (1983). *Introduction to Modern Information Retrieval*. New York, NY: McGraw-Hill.
- Schechter, S., Krishnan, M., & Smith, M. D. (1998). Using path profiles to predict HTTP requests. In *Proceedings of the seventh international conference on World Wide Web 7*. Brisbane, Australia.
- Schmettow, M. (2005). Towards a pattern based usability inspection method for industrial practitioners. In *Proceedings of the Workshop on Integrating Software Engineering and Usability Engineering*. Retrieved August 24, 2008, from

http://www.se-hci.org/bridging/interact2005/03_Schmettow_Towards_UPI.pdf

- Segall, M. H. (1990). *Human behavior in global perspective: an introduction to cross-cultural psychology*. New York, NY: Pergamon Press.
- Sheth, B., & Maes, P. (1993). Evolving agents for personalized information filtering. In *Proceedings of the 9th IEEE Conference on Artificial Intelligence for Applications*.
- Sheth, B. (1994). *Adaptive Agents for Information Processing*. SM Thesis, MIT.
- Sherman, C. (2000, July). Humans Do It Better:
  Inside the Open Directory Project. *Online*.
- Shneiderman, B., Nielsen, J., Butler, S., & Levi, M. (1998). *Is the web really different from everything else?* Paper presented at the Proceedings of CHI 98 summary conference: Human factors in computing systems, Los Angels, CA.
- Srikant, R., & Agrawal, R. (1995). Mining Generalized Association Rules. In *Proceedings of 21th International Conference on Very Large Data Bases (VLDB'95)*. Zurich, Switzerland.
- Song, H., Curran, E. P., & Sterritt, R. (2004). Multiple foci visualisation of large hierarchies with FlexTree. *Information Visualization*, *3*(1), 19-35.
- Spencer, R. (2000). The Streamlined Cognitive Walkthrough Method, Working Around Social Constraints Encountered in a Software Development Company. *CHI 2000*, *2*(1), 353-359.
- Spool, J., & Schroeder, W. (2001). Testing websites: Five users is nowhere near enough. In *CHI '01 extended abstracts on Human factors in computing systems* (pp. 285-236). Seattle, Washington.
- Stasko, J. T., Catrambone, R., Guzdial, M., & Mcdonald, K. (2000). An evaluation of space-filling information visualizations for depicting hierarchical structures. *International Journal of Human Computer Studies*, *53*(5), 663-694.
- Stasko, J. T., & Zhang, E. (2000). Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations. *INFOVIS 2000*. Salt Lake City.

- Stefani, A., & Strapparava, C. (1998). Personaliziong Access to Web Sites: The SiteIF Project. In *2nd Workshop on Adaptive Hypertext and Hypermedia (held in conjunction with HYPERTEXT '98)*. Pittsburgh, USA.
- Steinberg, S. G. (1996, May). Seek and Ye Shall Find (Maybe). *Wired*, (4.05). Retrieved August 24, 2008, from http://www.wired.com/wired/archive/4.05/indexweb.html?pg=1&topic=&topic_set=
- Sullivan, D. (1997, September 3). Yahoo: Delays Expected. *Search Engine Watch (SEW)*. Retrieved September 25, 2008, from http://searchenginewatch.com/showPage.html?page=2165541.
- Sullivan, D. (1997, September 3). Yahoo Special Report. *Search Engine Watch (SEW)*. Retrieved August 19, 2008, from http://searchenginewatch.com/showPage.html?page=2165591
- Sullivan, D. (1997, September 3). Yahoo Submissions Survey. *Search Engine Watch (SEW)*. Retrieved August 24, 2008, from http://searchenginewatch.com/showPage.html?page=2165571
- Sullivan, D. (1998, April 30). Directories take center stage. *Search Engine Watch (SEW)*. Retrieved August 19, 2008 from http://searchenginewatch.com/showPage.html?page=2166101
- Sullivan, D. (1999, May 4). Lycos Transforms Into Directory. *Search Engine Watch (SEW)*. Retrieved August 19, 2008, from http://searchenginewatch.com/showPage.html?page=2167171
- Sullivan, D. (2005, Jan 28). Search Engine Sizes. *Search Engine Watch (SEW)*. Retrieved August 19, 2008, from http://searchenginewatch.com/showPage.html?page=2156481

T

- Taylor, A. G. (1992). *Introduction to Cataloging and Classification* (8th ed.). Englewood, Colorado: Libraries Unlimited.

- Taylor, A.G. (1999). *The Organization of Information*. Englewood, Colorado: Libraries Unlimited.
- Taylor, A. G. (2000). *Wynar's Introduction to Cataloging and Classification*. Englewood, Colorado: Libraries Unlimited.
- Technology. *Alexa Company Info*. Retrieved October 1, 2008, from http://www.alexa.com/site/company/technology
- The History of Yahoo! - How It All Started... .*Yahoo! Media Relations*. Retrieved August 24, 2008, from http://docs.yahoo.com/info/misc/history.html
- Thomas, A. R. (1997). Bibliographical classification : The ideas and achievements of Henry E. Bliss. *Cataloging & classification quarterly* , *25*(1), 51-104.
- Thomson, L. (2005). A Standard Framework for Web Personalization. In *Proceedings of the 1st International Workshop on Innovations In Web Infrastructure (IWI 2005) co-located with the 14th World-Wide Web Conference (WWW2005)*. Japan.
- Tiamiyu, M. A. (1992). The relationships between source use and work complexity, decisionmaker discretion and activity duration in nigerian government ministries. *International Journal of Information Management*, *12*(2), 130-141.
- Turo, D. (1994). Hierarchical visualization with treemaps: making sense of pro basketball data. *Conference companion on Human factors in computing systems* (pp. 441-442). New York, NY: ACM Press.
- Tushman, M. L. (1978). Technical communication in R&D laboratories: The impact of project work characteristics. *Academy of Management Journal*, *21*(4), 624-645.
- Tuzhilin, A., & Adomavicius, G. (1999). Integrating user behavior and collaborative methods in recommender systems. In *CHI'99 Workshop Interacting with Recommender Systems*.
- Tzitzikas, Y., Spyratos, N., Constantopoulos, P., & Analyti, A. (2002, October).

Extended Faceted Taxonomies for Web Catalogs. ERCIM NEWS, (51).

U

- Uddin, M. N., & Janecek, P. (2007). Faceted classification in web information architecture, a framework for using semantic web tools. *The Electronic Library*, *25*(2), 219-233.
- Uddin, M. N., & Janecek, P. (2007). The implementation of faceted classification in web site searching and browsing. *Online Information Review*, *31*(2), 218-233.
- UK Internet User Monitor Survey. (2000). *Forrester Research*.

V

- Van Meteren, R., & Van Someren, M. (2000). Using Content-Based Filtering for Recommendation. In *ECML2000 Workshop*.
- Van Boskirk, S. and Li., C. (2001). Driving Customers, Not Just Site Traffic. *Forrester Research*. Retrieved August 19, 2008, from http://www.forrester.com/ER/Research/Brief/Excerpt/0,1317,12053,00.html
- Van de Ven, A. H., & Ferry, D. J. (1980). *Measuring and accessing organizations*. New York, NY: Wiley.
- Vickery, B. C. (1960). *Faceted Classification: A guide to the construction and use of special schemes*. London: Aslib.
- Vickery, B. C. (1966). *Faceted classification schemes*. New Brunswick, NJ: Rutgers State University.
- Virzi, R. A. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors*, *34*(4), 457-468.
- Vizine-Goetz, D. (1996). Using Library Classification Schemes for Internet Resources. In *Proceedings of the OCLC Internet Cataloging Colloquium*. San

Antonio, Texas.

- Vizine-Goetz, D. (2002). Using Library Classification Schemes for Internet Resources Revisited. *Journal of Internet Cataloging*, *5*(4), 5-18.
- Volokh, E. (2000). Personalization and Privacy. *Communications of the ACM*, *43*(8), 84-88.

W

- Walt, M. S. (1997). The role of classification in information retrieval on the Internet: some aspects of browsing lists in search engines. In *Knowledge Organisation for Information Retrieval: Proceedings of the 6th International Study Conference on the Classification Research* (pp. 32-35). London: University of London.
- Web directory. (2008, August 9). *Wikipedia, The Free Encyclopedia*. Retrieved August 19, 2008, from http://en.wikipedia.org/w/index.php?title=Web_directory&oldid=230820952
- Wharton, C., Rieman, J., Lewis, C., & Polson, P. (1994). The cognitive walkthrough method: A practitioner's guide. In *Usability Inspection Methods* (pp. 105-140). New York, NY: John Wiley & Sons.
- What Can I Measure? (n.d.). *Can Usability Be Measured? - Usability Basics*. Retrieved July 13, 2009, from http://www.usability.gov/basics/measured.html
- What's so great about the Dewey Decimal System? (2006, January 31). *The Straight Dope*. Retrieved August 24, 2008, from http://www.straightdope.com/columns/read/2238/whats-so-great-about-the-dewey-decimal-system
- Wheeler, D. (Ed.). (1996). *Software Inspection : An Industry Best Practice*. Los Alamitos, CA: IEEE Computer Society Press.
- Who We Are And What We Do. *ODP – Open Directory Project*. Retrieved August 19, 2008, from http://www.dmoz.org/help/geninfo.html

- Williamson, N. J. (1997). Knowledge structures and the Internet. In *Knowledge Organisation for Information Retrieval: Proceedings of the 6th International Study Conference on the Classification Research* (pp. 23-27). FID, The Hague.
- Wilson, T. D. (1996). On user studies and information needs. *Journal of Documentation*, *37*(1), 3-15.
- Wilson, T. D. (1999). Models in information behaviour research. *Journal of the American Society for Information Science*, *55*(3), 249-270.
- Wilson, T. D. (2000). Human information behavior. *Special Issue on Information Science Research*, *3*(2), 49-56.
- Wixon, D., Jones, S., Tse, L., & Casaday, G. (1994). Inspections and design reviews: Framework, history, and reflection. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods* (pp. 79-104). New York, NY: John Wiley & Sons.
- Wood, R. E., Mento, A. J., & Locke, E. A. (1987). Task complexity as a moderator of goal effects: a meta-analysis. *Journal of applied psychology*, *72*(3), 416-425.
- Woolrych, A., & Cockton, G. (2001). Why and when five test users aren't enough. In *Proceedings of IHM-HCI 2001* (Vol. 2, pp. 105-108). Toulouse. France.
- Wright, P. C., & Monk, A. (1991). A Cost-Effective Evaluation Method for Use by Designers. *International Journal of Man-Machine Studies*, *35*(6), 891-912.
- Wynar, B. S. & Taylor, A. G. (1992). *Introduction to cataloging and classification* (8th ed.). Libraries Unlimited.

Y

- Yahoo! Directory Guidelines for Submitting Descriptions. *Yahoo! Directory Help*. Retrieved August 24, 2008, from http://help.yahoo.com/l/us/yahoo/directory/ctd/ctd-02.html
- Yahoo! Ranked No. 1 Web Site Among Business Users in Ever PC Meter

Workplace Study. (1997, July 21). *Yahoo! Press Releases*. Retrieved August 19, 2008, from http://docs.yahoo.com/docs/pr/release106.html

- Yahoo! Ranked No.1 Navigational Network in April. (1997, June 16). *Yahoo! Press Releases*. Retrieved August 19, 2008, from http://docs.yahoo.com/docs/pr/release98.html
- Yahoo! Expands Audience; Ranked No. 1 in U.S. and Overseas. (1998, April 8). *Yahoo! Press Releases*. Retrieved August 19, 2008, from http://docs.yahoo.com/docs/pr/release169.html
- Young, R. M., Green, T. R. G., & Simon, T. (1989). Programmable user models for predictive evaluation of user interface designs. In *Proceedings of CHI'89 Conference on Human Factors in Computer Systems* (pp. 15-19). New York, NY: ACM Press.

Z

- Zakon, R. H. (2006). Hobbes' Internet Timeline v8.2. In *Zakon.org*. Retrieved August 19, 2008, from http://www.zakon.org/robert/internet/timeline/
- Zakon, R. H. (1997). *RFC2235: Hobbes' Internet Timeline*. IETF. Retrieved August 24, 2008, from http://www.ietf.org/rfc/rfc2235.txt
- Zhang, Z., Basili, V., & Shneiderman, B. (1999). Perspective-based Usability Inspection: An Empirical Validation of Efficacy. *Empirical Software Engineering*, *4*(1), 43-69.
- Zins, C. (2002). Models for classifying internet resources. *Knowledge Organization*, *29*(1), 20-28.