Research Paper

Juxtaposing the performance of Integrated and Separated Systems

Author information:

- C.P.A. Lemmens
  Information Science
  University of Utrecht
  e-mai: cpa@cpa-design.com

- Supervisor:
  Dr. M.C. Puerta Melguizo
  University of Utrecht

Table of Contents

Note that the Methodology, Results and Appendixes were cut-out. If you are interested in the full research you could contact me personally by email: cpa@cpa-design.com

Table of Figures

Abstract

Search and browse systems were developed to help us find information quickly. However, today's technology takes us a bit further and tries to understand our information goals. Understanding them should improve the search time and accuracy of the results. The systems that aim to understand the users' information goals and to integrate search and browse according to them, are what we decided to call Integrated Systems. The main concern of the study was to explore whether integrated systems perform better than the separated ones. Our hypothesis was that integrated systems will fulfil in the information goal with greater efficiency than separated systems. To prove this our subjects were asked to execute a series of tasks with separated and integrated systems. Unfortunately our results show that our hypothesis was not entirely correct. Systems probably will be more effective without integrating browse and search, we named such systems Dynamic Search Systems. These systems understand the user's information goal through understanding the semantic relations of keywords. It offers, based on these semantic relations, alternatives to users' searches and maintains a very simple interface. Within the conclusions the evidence is explained and a final advice given.

Introduction

The complexity of today's enterprises, the multiplicity of sources and the constant changes in hardware, software, threats and regulations, make absolutely critical to find the demanded information quickly. It is for this reason that the world's leading provider of research and analysis about global information technology industry, Gartner (http://www.gartner.com/DisplayDocument?ref=g_search&id=458917), published in 2004 a research written by K. Esteban that stated: "Selecting the proper search method or combination of methods, for a customer interaction can save an organization up to 40 percent in related maintenance costs. The key is knowing how to match methods with interactions." The purpose of the study at hand is to try to prove this statement by assessing users' performance using a system that integrates browse and search versus their performance with non integrated systems. We call systems that integrate browse and search *integrated systems* and those which do not, we call them *separated systems*. The research question of this study is: "which approach, the integration or separation of browse and search systems, performs best?". We refer to performance as the time it takes to return the appropriate results with the highest level of accuracy. In other words, the time it takes for the user to reach her information goal with the best possible results.

Before diving into the real investigation we first want to introduce its cause, the Accessibility Knowledge Base Project.

The Accessibility Knowledge Base Project

This project was part of C.P.A. Lemmens' internship which was conducted at the Accessibility foundation. He was concerned with the development and design of an interactive knowledge base, which includes a discussion board and high-end administrator/moderator facilities.

The knowledge base was designed to provide problem solving articles about tools and other products used by the impaired Dutch. The foundation's mission is to stimulate the accessibility of websites for all people, included disabled people. Consequently, the main challenge was to build a system that was fully accessible for disabled users that might suffer from disabilities such as visual, hearing, motor, and cognitive impairments. As a result of these expectations, the starting point was a review of the existing literature resulting in the full implementation of the WAI I, II and III recommendations[1].

The system includes special administrator/moderator features such as advanced article comprehension modules; image manipulation; advanced user management and article edit management systems; system configuration facilities, etc. Besides, the system includes as series of user-based-features as well, such as: browse and search systems, discussion facilities, and more. The system can be found at *http://www.tipentruc.nl* and is entirely written in ASP 3.0. It uses SQL algorithms to manipulate the MS Access 2003 database that has a capacity of 2GB storage space.

[1] http://www.w3.org/TR/WAI-WEBCONTENT

The browse ontology is quite simple (see figure 1). First users see the main folders; here not only the folder's name is provided but also a small description of its content and the number of articles inside. After selecting a folder the second layer within the ontology is reached; here the articles are presented and ordered by product name. The choice to order by product name was made since many users do own the product they want information on. The second line contains the article category and name to help users understand the meaning of the article quickly. After users select an article some related articles are presented below the actual article. Every page within the Knowledge Base also includes the option to search as can be seeing in the figure below.

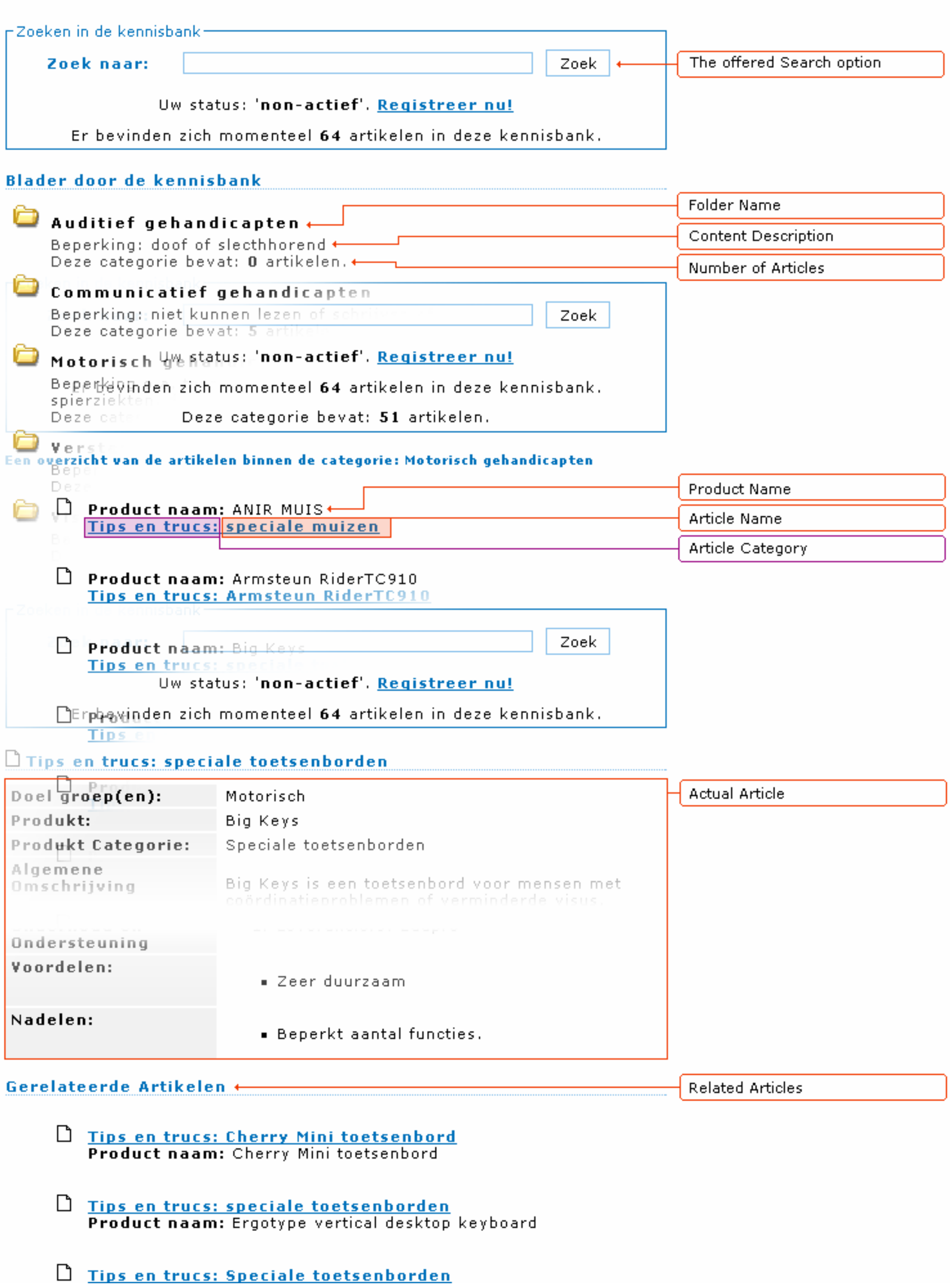

Figure I - The Search Unit and the Browse Ontology

We just gave you an inside on this project, now it's time we present its relations to this study, to be considered the second phase of the internship.

The relation

There were many influences that shaped this research and although the initial research question was about disability issues in addition to websites, this question had to be drastically modified. Circumstances as the unavailability of impaired subjects caused by broken promises of the Accessibility foundation, budget and time issues, etc. greatly affected our original research question. Below we present the original quote from my project manager that states that because of certain circumstances, the investigation was delayed and as a consequence impaired subjects became unavailable.

> "*Door omstandigheden heeft ons onderzoek (Kennisbank) enige vertraing opgelopen waardoor het onderzoek met respondenten in de problemen komt. Om het onderzoek te versnellen stel ik voor dat wij het onderzoek naar de mate van gebruiksvriendelijkheid van de Kennisbank voor gehandicapten op een andere manier invullen. De respondenten komen niet naar Accessibility maar krijgen de opdrachten thuis gestuurd. Zij voeren de opdrachten op hun eigen computer uit en beschrijven zelf alle mogelijke problemen die zij tegenkomen*"

Then we decided to focus on a more specific part of the system: the browse and search systems. By comparing the Accessibility Knowledge Base against other information search systems, we came to conclude that there are two main types of systems: separated and integrated systems. Briefly we define separated systems as systems that work completely independently from each other and which do not learn from previous users' behaviour. Integrated systems do learn from the previous user's browse and search behaviour and try to understand the user's information goal.

Despite the fact that the Accessibility Knowledge Base is a separated system we thought that the development of an integrated system would be much more efficient and satisfying because they try to understand the user's information goal. In the next sections we will explain the differences between separated and integrated systems in more detail and the logic of this research.

Browse and Search Re-Defined

Currently there are two main paradigms for finding information online: browse and search (Olston & Chi, 2003). Browse, by default, refers to the act of finding online information through quickly scanning a conceptual schema, or ontology. Many database driven websites, the focus of this investigation, support a special 'browse mode'. Within the browse mode one can flip through fields and records quickly. This means that the standard navigation unit within a website is excluded from the 'browse mode'. The example below (see figure II) will clarify this statement.

Figure II - An example of a browse mode

The "Navigation Panel" in blue, is considered to be excluded from the browse mode, which is surrounded by an orange border. It is this area that adapts to the users' choices. We have chosen the 'Communicatief gehandicapten' folder, and as you can see in figure III the ontology adapted itself but the Navigation Panel remained consistent.

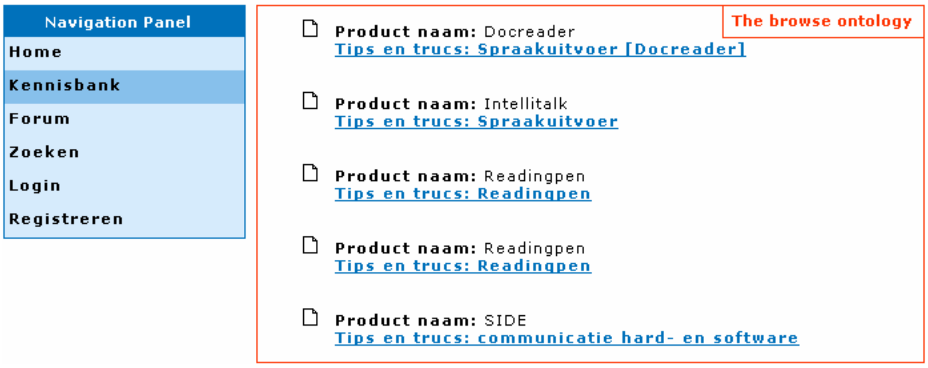

Figure III - An example of a browse mode

On the other hand, search is the act of finding information with a search engine. A search engine is a system that searches a database and gathers and reports information that contains or is related to the specified keywords. The execution of a search requires the user to specify her/his information goal into keywords. Normally the process is as follows: the user enters the keywords in the provided input box and presses the search button. The search engine will search the database and return results that match the keywords.

Both systems have their advantages and disadvantages, which we deal with next.

The complexity of browse and search systems

We have already mentioned the term "information goal" a number of times, but never really defined it, until now. The term information goal was first brought to our attention by Olston and Chi (2003) and they defined it as: some specific information users are seeking. However, we modified their definition since we distinguish between two kinds of information goals: the clear and the complex ones.

We defined an information goal as: a demand for a piece of specified information. Olston and Chi used the word 'seeking' which refers to a *fulfilment* whereas a 'demand' refers to *the need to fulfil*. Another small difference is the fact that we refer to 'specified information' instead of specific since a demand can occur in two forms: *defined or obscure*, and specific refers to an *defined-demand* only.

A clear information goal refers to a defined-demand whereas a complex information goal refers to an obscure-demand. Before we continue with this subject let us first point out the complexities of browse and search systems.

Search' strength is its ability to identify pages often in a reasonable time and especially when users have clear information goals. Browse on the other hand, seems to support users without a clear information goal and might even support them by crystallizing their information goal (Olston and Chi, 2003). These are clearly the advantages of the systems, but what about the disadvantages? There are times that the information goal does not lend itself to be transformed into keywords or that the users is not familiar with the website' terminology. In such times browse might be more helpful than search (Olston and Chi, 2003), but often both systems fail to fulfil the user's information goal.

There are researchers whom argued that a clear information goal results in the use of search systems (Choo et al., 1999, Jul & Furnas, 1997, Olston & Chi, 2003). However, according to Teevan et al. (2004) only about 35% of the users with a clear information goal actually use search first; more frequently they start by browsing. They argue that the reason for this has its roots in the fact that users often are not capable of transforming their information goals into meaningful keywords to conduct a search. Nevertheless, after reading the articles of Navarro-Prieto et al. (1999) and Hölscher & Strube (2000) we think that another reason for this could be the differences of expertise. Navarro-Prieto et al. (1999) identified cognitive strategies that are related to 'web searching'. They compared novice and expert 'web searchers', or in other words people with little or much experience in search. They found that expert searchers plan their search ahead based on their previous experience about the functionality of search engines, while novice searchers hardly plan at all and are rather driven by the external representations; the browse mode. This also implies that expert searchers will start their search for information by using the search engine first. In addition, Hölscher & Strube (2000) conducted two experimental studies and found that users with more web experience and knowledge domain find their information goal sooner. Based on their findings, we thought that it is not only important to have a clear information goal but also to have some knowledge about the domain in order to use search efficiently. Besides we expected that the more complex the information goal becomes the more the browse mode will be used. Within the discussion of our results we will get back at these expectations.

Although browsing provides a very simple mechanism for information scanning, it can be very time consuming and result in frustration. The main cause of this complexity lies in the fact that similar named ontologies may contain different types of pages (Gauch et al., 2003; Teevan et al., 2004). For example, a user may find that a topic belongs to the 'Thriller' category while another user may consider it to belong to the 'Detectives' category. However, the alternative; search, has its own problems. It has been estimated that about one half of all retrieved documents are irrelevant (Casasola, 1998) and the main reason for this poor result hides in the fact that words do often have multiple meanings and therefore are wrongly interpreted by the search engine (Krovetz, 1992).

That means that in both cases the meaning of words can be the cause of failing to fulfil in a user's information goal. This leads us to the necessity for a solution to resolve those situations.

The integrated system

We, hopefully, clarified that neither paradigm browse nor search alone is adequate enough for the more complex information goals (Olston and Chi, 2003), and that leads us to the solution which we call: *integrated systems*.

We define an integrated system as a system that adapts to, and understands the user's information goal based on previous user's browse and search behaviour. In other words, the system keeps track of what the user has been looking at, and what these pieces of information have in common based on their semantic relations. That information serves as for the creation of a short-term profile and to adapt the returned results, improving the accuracy of these results.

Much research has gone into user-understanding and how to store user preferences based on previous user behaviour (Truong et al. 2002; Alaniz et al. 2003). These user preferences are stored in so called profiles. We can distinguish between two kinds of profiles: the (explicit) long-term and the (implicit) short-term (Sugiyama et al., 2004). Long-term-profiles are created, and therefore also maintained, by the users. In contrast, short-term-profiles are based on navigation captures, identification of semantic relations between the browsed documents and frequency of the keywords used (Alaniz et al. 2003).

It is assumed that when a user's information goal changes, the long-term-profile will negatively affect the understanding of the user's information goal resulting in less accurate results (Pan, X. 2000; Pirolli et al, 2002). Prove for these assumptions can be found within the following example. In the earlier times, when Amazon was only a long-term-profile based system, there was a user who bought a book about pregnancy as a gift to his wife. Later when he returned he got a great many books about pregnancy recommended, and believe me, this was not his topic of interest (Chi, 2004).

The short-term-profiles on the other hand, are believed to increase the accuracy of results (Alaniz et al., 2003; Pirolli et al, 2002). Combining short-term and long-term profiles by adapting the long-term-profiles into ones that possibly can give a greater understanding of the user's information goal is an option; however, this option is often omitted because of its complications, e.g. the need to understand the users' information goals in order to create effective combinations, only a difficult and time consuming study will deliver the demanded results and the fact that regular content changes will affect the users' information goals and cause the need for new studies.

Although the Accessibility Knowledge Base is a separated system it does have a function to store a long-term profile. Registration allows users to identify their disability and based on that information we offer an option to browse or search based on their disability. A disability remains consistent and does not change, at least not as often as an information goal. It is for this reason that this function will prove effective, however, this is not to be considered as an example of a tool that belongs to an integrated system. The system does not adapt the user's information goal, neither does it learn from the previous user behaviour and nor are the browse and search systems integrated.

An integrated system put to work

The websites of Amazon are integrated systems, we focus at Amazon.co.uk, built in Perl/CGI, ran at an UNIX platform and put to work in combination with external programs written in C/C++ and Java to establish its performance. Amazon system's short-term-profiles are based the previous browse and search behaviour of the user. We assume that the recommended items are also a result of clever short-term-profile combinations with database managed meta-information. This assumption is based on the fact that after you have been actively looking for a specific item you will find that it updates its recommended items as well[2].

Related work

The fields of computer and information sciences have paid little attention to the investigation of integrated versus separated systems. However, there have been a few studies investigating the topics related to the integration of browse and search systems (Olston & Chi., 2003; Pirolli et al., 2000 and 2002; Choo & Marton 1999; Sugiyama et al., 2004). Nonetheless, most of these studies did not explore the effects of the systems' performance differences and that makes this study unique and exploratory.

In the first experiment of Hölscher and Strube (2000) 12 established internet experts were interviewed about their normal search strategies. After that, they performed a series of realistic search tasks on the World Wide Web. Within the second experiment two types of potentially relevant types of knowledge were compared: Web experience and domain-specific background knowledge. Subjects were asked to perform a series of search tasks in an economics related domain. This resulted in the following conclusion: web experience and domain-knowledge combined increase the ability to successfully perform a search. But, probably one of the best-known studies in the area of integrated versus separated systems is the one of Olston and Chi (2003).

They describe the two predominant paradigms for finding information the web: browsing and keyword searching. They argue that while both paradigms exhibit complementary advantages, neither is adequate alone for complex information goals that lend themselves partially to browsing and partially to searching. According to them, this is because complex information goals are a result of several criteria and often with some criteria in different search categories. For instance suppose a user is looking for a 'fast digital copier', in that case, search might be more appropriate for finding the 'digital copiers' (the user could simply use the keywords 'digital' and 'copiers' to find the relevant information). Browsing is possibly more suitable to identify the 'fast' copiers; since it is a relative criterion and is difficult to represent using keywords search. According to the authors, it is important to integrate browsing and searching into a single interface and propose the approach called ScentTrials to reach this goal. Based on short-term profiles, ScentTrails highlights hyperlinks to indicate path to search results

[2] Note that Amazon was not willing to give us any details about its system; it is for this reason that we had to use other sources which we are not allowed to mention.

and users can interpolate between searching and browsing to find appropriate information (Chi, 2004). In a preliminary evaluation of the prototype, they found that ScentTrails enabled users to find information more quickly than by either searching or browsing alone.

A variety of methods have been used to study information seeking and retrieval. Laboratory studies for instance have allowed researches to conduct controlled studies and examine users' thought processes during the search by having them think aloud as they search (Muramatsu and Pratt, 2001; Hölscher and Strube, 2000). Muramatsu and Pratt (2001) found that users have difficulties understanding the operation of query transformations without additional assistance. They propose to use 'Transparent Queries' that make the normally opaque transformations and processing of a user's query more transparent by illustrating the transformations in a visually annotated presentation of the query (Muramatsu et al., 2001). They conducted a user study to gain a better understanding of users' knowledge and reactions to the operation of several query transformations that web search engines automatically employ. The results of this study suggest that users do indeed have difficulties understanding the operation of query transformations without additional assistance. Although transparency is helpful and valuable, interfaces that allow direct control of query transformation by its users might ultimately be more helpful. It can be helpful because it visualizes to the user what actions are taken to return more effective results. An example to explain this is given next. The query "to be or not to be" could return an actual query as "be", all other words were removed because they are too common and as you can imagine this will have a major affect on the returned results. Therefore we think that an integrated system is not only supposed to learn from previous user behaviour but also to offer a sort of transparency, in other words: relations or recommendations should be visualized. An example of how to implement query transparency could be: Your actual query was, [query], and became [query], if you wish to search for an exact query use quotes: ["query"]. The idea of transparency should definitely be considered in integrated systems, but transparency alone is by far not good enough since it only visualizes the query transformation but does not offer any semantic relations or whatsoever.

Sugiyama, Hatano and Yoshikawa (2004) conducted a study that actually introduces semantic relations. Their study is based on the fact that separated systems only return static results even when users may change their information goals from one moment to another. Therefore they proposed several approaches to adapt search results according to each user's information goal without any user effort, and verified the effectiveness of these approaches. The proposed approaches to verify effectiveness were: (1) users explicit feedback about their current information goal and two different implicit approaches: (2) user profiles based on pure browsing history and (3) user profiles based on the modified collaborative filtering algorithm they developed which includes semantic relations. While users did have to provide feedback explicitly in (1), they did not have to provide any effort in the proposed methods (2) and (3) since their system implicitly captured changes in user's preference. The experiments were implemented using Perl. They used 50 modified query topics that were employed as test topics. After the subjects submitted these queries to Google, their system reordered the search results according to each user's profile constructed with help of their algorithms. They observed the browsing history of 30 subjects for 30 days. They found that search systems that adapt to each user's preferences can be achieved by constructing short-term user profiles based on modified collaborative filtering with detailed analysis of users' browsing history in one day. Their approach allowed each user to perform a fine-grained search, which is not performed in typical search engines, by capturing changes in each user's preferences and it was this approach that proved most effective.

Conclusions

Which approach, the integration or separation of browse and search systems, performs best? That is the question we will answer in this section. The answer we expected to find was that the integrated system would perform better; however, the investigation proved otherwise.

As we saw within the results most terminations happened during the complex tasks and especially Amazon suffered here. In addition it was again Amazon that demanded most time; correct and incorrect, for both browse and search. However it was Accessibility that demanded most steps and not Amazon. When we look at the keyword averages we see that once again it was Accessibility that showed the greatest use of keywords and not Amazon. Nevertheless the final average grades given by the subjects showed nearly no differences between Accessibility and Amazon. Then we have Gartner, which stood-out all the way. It proved to perform better within the complex tasks as opposed to the clear tasks considered time, steps, and grades. It was much faster and more efficient than its opponents again considered the complex tasks only. The grades Gartner received on the complex tasks were indeed higher than its opponents but lower considered the clear tasks, and so was its performance. This somewhat radical difference might be due to the fact that the clear tasks for Gartner were found harder to solve as the use of keywords indicated. This statement was supported by the questionnaires, where the subjects pointed out that these tasks was indeed harder to solve.

When we look back at the grades we see that the differences between the overall averages are very small, and it is that Amazon proved to be best performing, however, the differences are found minimal. Nevertheless our data demonstrates that it was Gartner that proved to be more efficient especially considered the complex tasks which was our main focus. Since it seems that our data is not really to resolve the doubt about which system performed best we took a closer look at the actual feedback we received from the questionnaires.

Subjects were asked if the given tasks were clear enough. Their feedback made clear that the 4th task in both versions was found more difficult. However, only within version B there were 4 out of 7 subjects whom indicated to experience trouble comprehending the 5th task, These 4 subjects were not capable of solving the task either and were the cause of termination as visible within graph 1. We took a closer look at the remaining three subjects whom successfully solved the concerned task. Two of them were females and one male, he was a student Law. We think that the other subjects did not read as carefully as those three did but we cannot really support this statement otherwise than the fact that women in general show a greater understanding of the material they processed (Anderson, 2005).

Also we found that people who studied Computer Science, Information Science or Artificial Intelligence performed much worse than the average subject. Four subjects out of the five terminations were students active in one of these fields. Besides the in total 7 subjects active in one of these fields all had more then 5 years of internet and computer experience; nonetheless they somehow managed take consume most time. According to Hölscher & Strube (2000) and Navarro-Prieto et al. (1999) we are supposed to categorize these subjects into the category of 'expert-searchers'. But according to them it were the 'expert-searchers' that should perform best, which was not the case within our investigation. We therefore could argue that their definition of 'expert-searchers' is incorrect or that they, or we, came to wrong conclusions. Therefore we will leave this point open for further discussion. It is clear that this issue requires further investigation.

There were no problems with the systems' interfaces, due to the familiarity tasks. However, regardless we set the start screen at the search- or browse-mode users tended to start by search. There were some exceptions but most often caused by the misunderstanding of their freedom to choose between browse and search. This means that users tend to search and not to browse which contradicts with the results Teevan et al. (2004) found. They argued that only about 35% of the users with a clear information goal actually use search first, we came to conclude that this percentage is way to low since it turned out to be 81,08% in our investigation. Besides we expected that the more complex the information goal becomes the more the browse mode will be used, this expectation is supported by our data as is visible in graphs 4 and 11.

We admit that comprehending difficulties had an affect on the results but we do not think that these affects were great enough to cause such large differences between Gartner and Amazon. However, considered all evidence our conclusion remains; Gartner performed best despite the small differences between Gartner and Amazon.

In order to explain this statement we looked at a number of other concerns. We found was that users only switched to the browse mode when they had reached the level of total desperateness. Next we consider that Google was found, by all subjects, the preferred search engine. During the experiments subjects often referred to the '*did you mean'* function Google offers, except during the tasks where Gartner's system was used since it already offers this function. Amazon does have a certain '*did you mean'* function but lacks to show this which sometime let to frustration by subjects. Nevertheless as soon as subjects found out that Amazon was actually able to understand their information goal it let to great excitement. These facts let us to believe that main function of integrated systems is certainly to be considered. However, in order to put the main function or *comprehensive module* to work one needs to realize that this module should be capable of understanding semantic relations between data; independent of the terminology used by the user and website. This means that the comprehensive module should be able to understand the meaning of words. We believe that a good comprehending module could avoid the user to reach the level of total desperateness and therefore the need to browse. We think that the philosophy of integrated systems is almost correct except for one small detail; browse should not be integrated.

Dynamic Search Systems

Which approach will perform best? The answer is simple: Dynamic Search Systems.

A Dynamic Search System can be defined as a system that understands the user's information goal through understanding the semantic relations of keywords. It offers, based on these semantic relations, alternatives to users' searches and its interface should be clear, simple and include query transparency. A Google like interface is recommendable since nearly all users will be familiar with it. We call this kind of systems: Dynamic Search Systems.

Future Work

In future work, we plan to conduct experiments with a greater number of subjects and attempt to improve our investigation design since we now need to include a Dynamic Search System and we now can eliminate the system with a small database, since this variable seemed not to matter much.

References

Alaniz, A., Truong, K.N., Antonio, J. (2003) *Automatically Sharing Web Experiences through a Hyperdocument Recommender System.*
Retrieved from the World Wide Web at the 25th of June 2005.
www.ht03.org/papers/pdfs/6.pdf

Anderson, J.R. (2005) *Cognitive Psychology and its Implications.*
Worth Publishers; New York.

Bates, M.J. (1989) *Where should the person stop and the information search start?*
Retrieved from the World Wide Web at the: 14th of Mars 2005.
http://www.gseis.ucla.edu/faculty/bates/searchinterface.pdf

BBc News. (2004) *Google rolls out corporate search*
Wednesday, 20 October, 2004, 13:57 GMT 14:57 UK
Retrieved from the World Wide Web at the: 10th of May 2005.
http://news.bbc.co.uk/1/hi/business/3759878.stm

Berners-Lee, T., Hendler, J., and Lassila, O. (2001) *The Semantic Web.*
In Scientific American, 284 (5), May, 2001, pp 34-43.
http://www.scientificamerican.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&catID=2

Belkin, N.J., Brooks, H.M., Oddy, R.N. (1983) *ASK for information retrieval.*
Journal of Documentation 38:61-71.

Casasola, E. (1998) *ProFusion Personal Assistant: An Agent for Personalized Information* Filtering on the WWW.
The University of Kansas; Master's thesis

Chi, E.H. (2004) *Transient User Profiling.*
Retrieved from the World Wide Web at the: 14th March 2005.
http://www2.parc.com/istl/groups/uir/pubs/items/UIR-2004-05-Chi-TransientUserProfiling.pdf

Choo, C.W. and Marton, C. (2003) *Information seeking on the Web by women in IT* professions. Internet Research, 13(4), 267-280.
Retrieved from the World Wide Web at the: 14th March 2005.
http://choo.fis.utoronto.ca/FIS/ResPub/InternetR/default.html

Choo, C., Wei, D., Brian and Turnbull, D. (1999) *Information seeking on the web — an integrated model of browsing and searching.*
In Larry Woods (Ed.) ASIS '99: Proceedings of the 62nd ASIS Annual Meeting. (Vol. 36) (pp. 3-16).
Medford: Information Today.
Retrieved from the World Wide Web at the: 14th March 2005.
http://www.firstmonday.dk/issues/issue5_2/choo

Dervin, B., and Nilan, M. (1996) *Information needs and uses.*
In: Annual review of Information Science and Technology, vol. 21. Ed. M.E. Williams.
New York: Knowledge Industry Publications, 3-33.

Esteban, K., (2004) Select the Right Search Method for the Job.
Retrieved from the World Wide Web at the: 20th July 2005.
Gartner.com - http://www.gartner.com/DisplayDocument?ref=g_search&id=458917

Gauch, S., Chaffee, J., and Pretschner, A. (2003) *Ontology-Based Personalized Search and Browsing.*
Retrieved from the World Wide Web at the: 3rd of May 2005.
http://portal.acm.org/citation.cfm?id=1016421
http://www.ittc.ku.edu/~sgauch/selectedpapers/WIAS2003.pdf

Hölscher, C., and Strube, G. (2000) *Web Search Behavior of Internet Experts and Newbies.*
Computer Networks, Vol. 33 no 1-6, 337-346.
Retrieved from the World Wide Web at the: 14th of Mars 2005.
http://www9.org/w9cdrom/81/81.html

Jul, S. and Furnas, G.W. (1997) Navigation in Electronic Worlds – Workshop Report.
Retrieved from the World Wide Web at the: 14th March 2005.
http://www.si.umich.edu/~furnas/Papers/Nav97_Report.pdf

Krovetz, R., and Croft.B.W. (1992) Lexical Ambiguity and Information Retrieval.
ACM Transactions on Information Systems, 10(2), April 1992, pp. 115-141.

Kutulthau, C. (1993) *Seeking meaning. A process approach to library and information services.*
Norwood, N.J. Ablex

Manber, U., Smith, M., and Gopal, B. (1997) *WebGlimpse – Combining Browsing and Searching.*
Retrieved from the World Wide Web at the: 14th March 2005.
http://webglimpse.net/pubs/webglimpse.pdf

Marchionini, G.M. (1995) *Information Seeking in Electronic Environments.*
Cambridge: Cambridge University Press.

Muramatsu, J., and Pratt, W. (2001) *Transparent Queries: Investigating Users' Mental Models of Search Engines.*
Retrieved from the World Wide Web at the: 14th March 2005.
http://www.ics.uci.edu/~pratt/pubs/SIGIR-TQ.pdf

Navarro-Prieto, R., Scaife, M., and Rogers, Y. (1999) *Cognitive Strategies in Web Searching.*
Proceedings of the 5th Conference on Human Factors and the Web, June 1999.
Retrieved from the World Wide Web at the 3rd of May 2005.
http://zing.ncsl.nist.gov/hfweb/proceedings/navarro-prieto/

Olston, C., and Chi, E.H. (2003) *ScentTrails: Integrating Browsing and Searching on the Web.*
Retrieved from the World Wide Web at: 14 March 2005.
http://www-db.stanford.edu/~olston/publications/scenttrails.pdf

Otter, M., and Johnson, H. (2000) *Lost in hyperspace: metrics and mental models.*
Retrieved from the World Wide Web at the: 14th March 2005.
http://www.cs.odu.edu/~jbollen/IR04/readings/otter2000.pdf

Pan, X. (2000) *An Experimental Adaptive Expert System.*
Retrieved from the World Wide Web at the 14th of May 2005.
www.cadrc.calpoly.edu/pdf/adaptive_expert.pdf

Pirolli, P., and Card, S.K. (1999) *Information Foraging.*
Retrieved from the World Wide Web at the: 14th March 2005.
http://www2.parc.com/istl/projects/uir/pubs/items/UIR-1999-05-Pirolli-Report-InfoForaging.pdf

Pirolli, P., Fu, W-T., Reeder, R., and Card, S.K. (2002) *A User-Tracing Architecture for Modelling Interaction with the World Wide Web.*
Retrieved from the World Wide Web at the: 14th March 2005.
http://www2.parc.com/istl/projects/uir/pubs/items/UIR-2002-07-Pirolli-AVI-UserTrace.pdf

Pirolli, P., Card, S.K., and van der Wege, M.M. (2000) *The Effect of Information Scent on Searching Information Visualizations of Large Tree Structures.*
Retrieved from the World Wide Web at the: 14th March 2005.
http://www2.parc.com/istl/projects/uir/pubs/items/UIR-2000-04-Pirolli-AVI2000-InfoScentAndHBSearch.pdf

Pollock, A., and Hockley, A. (1997) *What's wrong with internet searching. D-Lib Magazine.*
Retrieved from the World Wide Web at the: 1st May 2005.
http://www.dlib.org/dlib/march97/bt/03pollock.html

Schatsky, D., Berk, M., Card, D., Daniels, D., Blank, J., and Levy, M. (2003) *Web site search: understanding and influencing customer behaviour.*
Jupitermedia Corporation; Jupiter Research, Vol. 3.

Sugiyama, K., Hatano, K., and Yoshikawa, M. (2004) *Adaptive Web Search Based on User Profile Constructed without Any Effort from Users.*
New York; Retrieved from the World Wide Web at the: 3rd of May 2005
www.www2004.org/proceedings/docs/1p675.pdf

Teevan, J., Alvarardo, C., Ackerman, M.S., Karger D.R. (2004) *The Perfect Search Engine Is Not Enough: A Study of Orienteering Behavior in Directed Search.*
Retrieved from the World Wide Web at the: 14th March 2005.
http://haystack.lcs.mit.edu/papers/chi2004-perfectse.pdf

Truong, K.N., Abowd, G.D., da Graca Pimentel, M. (2002) *Vicariously Sharing Captured Web Experiences through an Automated Recommendation System*
Retrieved from the World Wide Web at the 29th of June 2005.
http://smartech.gatech.edu:8282/dspace/bitstream/1853/3312/1/02-26.pdf

Webopedia.com. (2005) *Browse.*
Retrieved from the World Wide Web at the: 26th of June 2005.
http://webopedia.com/TERM/b/browse.html

Wikipedia.org. (2005) *Search engine; Conceptual Schema; Ontology*
Retrieved from the World Wide Web at the: 12th of May 2005.
http://en.wikipedia.org/wiki/Search_engine
http://en.wikipedia.org/wiki/Conceptual_schema
http://en.wikipedia.org/wiki/Ontology_%28computer_science%29