

Internet Search and Navigation Strategies Used by Experts and Beginners

Carles Monereo, Marta Fuentes and Sònia Sànchez

Departament de Psicologia Evolutiva, Universitat Autònoma de Barcelona

carles.monereo@blues.uab.es

Abstract

This paper presents three pilot studies whose purpose was to analyse the way different users (experts and beginners) go about searching for specific information, both on the Internet in general and in a thematic web site in particular. It discusses the role of prior knowledge in achieving the search objectives, and highlights the differences in task completion due to the use of approaches of varying levels of strategic development.

Introduction

Nobody today denies the importance of the Internet as a powerful communication and information tool. The Net is a vast store of information that is steadily taking over from other means of communication. The number of users, information providers, and services in all areas of culture, consumption and leisure seems to be growing non-stop, and at a rate that would have been unthinkable only a few years ago. The impact on society has also gradually increased, to the point where the ability to find certain data on the Internet data so immediate or so specific that we could not find them anywhere else has become indispensable.

However, this is not an ability that comes naturally. There is evidence to suggest that all too often the search is routine, uncritical, unselective and unguided, based on trial and error. This would explain why so many cybernauts get "shipwrecked" on the Net.

In today's knowledge society, acquiring the skills and abilities to find information will increasingly become a vital necessity, and the failure to acquire them will define a new type of illiteracy and alienation from society.

This article deals with the strategies that users adopt in their search for information on the Internet and how they respond to the task of achieving a particular navigational goal. It also considers whether prior knowledge in any way affects the approach adopted by different users when faced with a search task.

1. Searching for Information on the Internet. Possibilities and Types

When looking for information on the Internet, there are various features of the Net that we need to take into account:

- The Internet is a plural, heterogeneous channel in which different means of electronic production, publishing and diffusion (WWW, e-mail, discussion groups, chat groups, etc.) interact. Each has its own set of rules, and they all relate to one another in many different ways, allowing for a great variety of hybrid forms.
- It allows for an exponential proliferation of information in multiple formats.
- This information is permanently available and can be changed and updated on line.
- Locating and, above all, evaluating this information often requires a degree of specialization on the part of the user.
- The information very often comes from unknown sources and, above all, is potentially unreliable.
- The information has not been systematically organized, and there is a clear bias in favour of certain subjects and in detriment of others.

Given the huge amount of information available and the fact that the Internet is not centralized, searching for information would often be a chaotic process if it were not for the fact that the Web itself offers various search tools or applications that help us to look for a particular document or topic.

It is essential to know from the outset exactly what kind of information we are looking for, because this knowledge will determine where we should look for it and which search tool we should use. The table below gives some basic guidelines for choosing the most appropriate search tool.

What are we looking for?	Where can we find it?
Information on a general subject	Directory
Information on a local subject	Local search engine
Information on a very specific subject	Search or metasearch engine
Personal web pages	Search engines
Web pages of organizations, institutions, companies, ...	Directories

E-mail addresses	Mail search engine
Books	Library catalogues ISBN Library web sites
Images	Directories

The information search services on the Internet can be classified in three groups: directories (collections of links to Internet addresses organized by subject matter or by geographical areas); search engines (systems that use a software robot or program to automatically search the Web for sites, index them, and enter them in a database); and metasearch tools (multiple search engines, programs that make it possible to search through various search engines at once).

2. Three Pilot Studies on Information Search Strategies

We shall now describe three pilot studies whose common goal is to analyse the variables that come into play in a web search, and the strategies for information search and Internet navigation

that can most contribute to the effectiveness of the search.

The first is a descriptive study which aimed to determine how university students use the Internet, particularly with regard to documentary resources. The second study compares the skills and cognitive and metacognitive strategies used by experts and non-experts in searching for information on the Web. The third study was designed to assess the importance of prior knowledge of a specific discipline when searching for information in a hypertextual environment.

2.1 *What use do university students make of information search via the Internet?*

We first conducted a descriptive study, using a specially designed survey, to find out what use the students at two Catalan universities one presence-based (Universitat Ram3n Llull) and the other virtual (Universitat Oberta de Catalunya, the Catalan Open University) make of the telematic network. Specifically, we wanted to establish which documentary search resources were the most commonly used; we then intended to compare them. Below we summarize the survey results that are most relevant to the subject of this article.

– No significant differences were found between the two universities, neither with regard to the most commonly used services (electronic mail, file transfer, and documentary sources, in that order), nor with regard to the type of search conducted: for a particular piece of information, for a

particular document, or for general information on a particular subject. The rest of the services (discussion groups, conferences, newsgroups, forums, chat, e-commerce, etc.) were used only very occasionally.

– Although 94% of the respondents considered searching for information on the Internet motivating, none of them gave this as his/her preferred source of information. Eighty-three percent continue to use print media as their main source of information; 12% use print media, CD-ROMs and the Internet to an equal extent; and 5% go straight to an expert for the information they need. For the respondents, what makes the Internet most useful as a source of information is:

- its speed and immediacy.
- the possibility of accessing a very wide range of information.
- the fact that it is a challenge, and finding information is gratifying (like a game).
- its convenience; users do not have to move, and do not have to consult a range of sources.
- the possibility of editing and manipulating the material found.

Six percent of respondents who did not consider the Internet a useful medium argued that it was slow and that having English as the main language of communication was an obstacle.

– The respondents showed a certain ignorance of search procedures and conditions. Those who said they did not embark on a search unless they had a known address also identified ignorance of search procedures (techniques, skills) as the main obstacle when it came to searching for information on the Internet.

- Sixty percent of the respondents tended to start their searches using the best-known search engines. They rarely mentioned the target of the search, the subject matter, or their own prior knowledge of the subject as important factors to be considered before starting the search. There is a strong tendency always to use the same search tools, regardless of the target of the search.
- The most commonly used instruments are search engines, followed by directories. Very few of the respondents regularly use metasearch tools; indeed, most are unaware of their existence or of their possibilities. The most commonly used worldwide search engines are Altavista, Yahoo, Lycos and Web Crawler. The most commonly used Spanish search engines are Olé, Dónde and the Spanish version of Altavista. Among Catalan search engines used are Vilaweb and Cercat.

– For most of the respondents, the first step is to enter the keyword for their search in the search engine. The purpose and the use of any information they may find and the time they have available do not usually condition their search behaviour. Indeed, 50% do not plan their search; 32% define the area of their search using very general terms; and only 18% do so using specific keywords.

– Predictably, those who have most difficulty in finding the information they are looking for are the ones who display the greatest lack of search resources, the greatest tendency always to use the same tools, the greatest difficulty with English, and the least ability to adapt their search procedure to the search objectives.

- When a particular Internet search proves fruitless or yields unsatisfactory results, 49% of the respondents tend to start another search using alternative search terms; 20% abandon the search; 1% seek outside help; and only 9% review the steps taken so far.
- In contrast, if the search yields a large amount of information, 39% tend to look for the information using another search engine; 27% select some of the documents obtained, applying fairly random criteria (the first in the list, those that include certain topics, etc.); 14% leave it at that; 11% restrict the search to more specific fields; and 9% review the steps taken so far in an effort to refine the search.
- Lastly, a majority of the respondents have great faith in the credibility of the information found.

In summary, the university students in our sample, regardless of whether they are engaged in presence-based or distance education, lack the necessary skills and strategies for searching for, selecting and using the information that the Internet has to offer. These shortcomings suggest that students need training in the use of appropriate search strategies. But what are the most appropriate strategies? To help answer this question, we conducted a second study.

2.2 What Internet search skills and strategies do experts use?

In order to ascertain what search strategies could be considered good models for training students we carried out a qualitative study with four Internet users, who were selected on the grounds of their extensive professional experience in its use. Basically, we asked them to describe in as much detail as possible how they looked for information on the Internet.

To do this we set two different search tasks, one more complex than the other (1: find the full text of main fairytales of the Brothers Grimm, and 2: find two versions of one particular tale). We asked the test subjects to describe aloud, step by step, how they would go about it (thinking aloud). We recorded these descriptions on audio cassette and subsequently made transcriptions in order to be able to analyse and compare the different ways of tackling the task. The following results emerged:

The first subject, who adopted what we could consider a strategic approach, explained the reasons for each of her actions, how she established an initial hypothesis regarding the content, why she used particular search tools rather than others, particular keywords rather than others, how she used her prior knowledge to interpret what she found, how she anticipated a search before actually doing it, what difficulties she encountered and how she solved them, etc.

After each step in the search process, she assessed whether she had succeeded in achieving the objective and reviewed the path she had taken, pointing out any changes that should be made.

The second subject adopted a fairly similar approach to the first one, but she had one major problem: her insufficient knowledge of English. As a result, her search was more limited to certain search procedures and was therefore less effective.

The third subject was much less strategic than the first; her actions were based on trial and error and improvisation. The only review she conducted consisted in acknowledging that what she found matched what she had requested. Not once did she consider what she was doing or why, nor whether she should change her approach in view of the difficulties encountered.

The fourth subject preferred first to look for basic information on the subject in an encyclopaedia and then use more accurate keywords.

A first conclusion of this pilot study is that a "strategic searcher" (one who plans, monitors, evaluates and steers the search process) follows a different cognitive search process from a "non-strategic searcher".

As can be inferred, this classification into "strategic" and "non-strategic" is partly independent of whether the person is a regular Internet user or not. We can therefore distinguish between non-strategic experts (if they have to act outside their particular area of expertise they find it difficult to target their search process) and strategic experts (even when they have to act outside their area of expertise they are able to devise alternative methods and find what they are looking for).

The searchers we have described as "non-strategic" barely stop to think or plan and decide what search tool to use. They start their search with the first tool that comes to mind and tend to jump from one step to another without any prior plan to guide their choice or the order in which they do things (the third subject used first a Spanish search engine, then a Catalan one, then a worldwide one, and then another Spanish search engine, quite arbitrarily). When they do not find what they are looking for or do not know where else to look, they give up (the third subject would have given up much sooner if it had not occurred to her to use a metasearch tool). If they review their actions at all, the review is limited to checking for possible spelling mistakes. They rarely question the approach they have adopted, their choice of search engine, the way they have organized the search, etc. Their monitoring activity is based on the simple acknowledgement of a match between the results and their request, while ensuring that the information they find is related to the subject of their search.

It has to be said, however, that even a search conducted in this manner may find the desired information, particularly if the searcher is familiar with the subject or with the different search tools and ways of searching. But this kind of uncontrolled search is unlikely to be effective in resolving complex and specific tasks that involve reorganizing knowledge. Some tasks involve deciding what is most important, paraphrasing the initial request or subject, using certain operators (AND, OR, etc.), finding analogies or synonyms, thinking of other things to include, etc. To do all this, it is essential to plan and reorganize the search, i.e. to go about it in a more reflexive, more strategic manner.

Searching strategically involves shifting attention constantly to and from between what we are looking for and how we go about looking for it. A strategic searcher adapts her behaviour to the objectives of the search: she takes the most appropriate decisions in the light of the information found on the web pages consulted, her own personal limitations, the circumstances of the search and the terms of the task; and she reviews progress at every stage of the search process, not just at the end.

In our study we observed how the more strategic searchers adapted to the demands of the situation by successively reorganizing and redirecting their actions. They reflected on what they were doing, constantly monitoring and reviewing their actions and the results they obtained; and when they detected a difficulty they acted immediately to adjust their action to the objectives of the search (adding or eliminating a keyword, switching to a different search engine, refining the search, setting geographical or chronological limits, using upper and lower case, checking spelling, using conditions such as and, not, or, etc.).

For this kind of searcher, strategies for monitoring and guiding one's own cognitive activity play an important role, given that deciding how to proceed at any given moment and controlling one's own decision-making process to achieve the desired objectives become fundamental. All this activity conscious, intentional, complex and flexible is more often recursive than linear.

As we have seen, a knowledge of the subject matter to be researched does not, on the face of it, seem to be the only determining factor for success in a search. To establish the role played by the searcher's knowledge of a particular subject area and of web search tools, we designed a third study.

2.3. What influence does the searcher's knowledge of the specific information content have on the search process?

Our objective in this last study was to investigate and analyse the variables that determine the effectiveness of an information search. Specifically, we wanted to analyse the impact of two variables: the searcher's knowledge of the subject to be researched; and the searcher's knowledge of the computer environment in which the search is carried out. In other words, we wanted to know how four different types of people a person who is an expert in the subject but knows nothing about computers (Es/Nc), a person who is an expert in computers but knows nothing about the subject (Ns/Ec), a person who is a novice in both the subject and computers (Ns/Nc), and one who is an expert in both (Es/Ec) handled the task in a hypermedia environment, and what cognitive skills each would have to mobilize.

Our starting hypothesis was that when tackling a complex information search task, adopting a more strategic approach would be the key success factor, and that this factor would intersect at various points with the other factors (subject competence or computer competence).

We proceeded as follows: first, we gave various people a questionnaire in order to select four, who are the ones we used as our sample. The selection was made on the basis of the individuals' knowledge of the subject and knowledge of computers.

SUBJECT A	<i>EXPERT IN THEMATIC CONTENT Ec</i>	<i>EXPERT IN COMPUTER CONTENT Ei</i>	Graduate in Sociology
SUBJECT B	<i>NOVICE IN THEMATIC CONTENT Nc</i>	<i>EXPERT IN COMPUTER CONTENT Ei</i>	3rd-year Computer Science student
SUBJECT C	<i>EXPERT IN THEMATIC CONTENT Ec</i>	<i>NOVICE IN COMPUTER CONTENT Ni</i>	Graduate in Educational Psychology
SUBJECT D	<i>NOVICE IN THEMATIC CONTENT Nc</i>	<i>NOVICE IN COMPUTER CONTENT Ni</i>	4th-year Psychology student

The web site chosen for the task was: "Aprendre a aprendre" ("Learning to Learn") (<http://www.xtec.es/~cdorado>). There were various reasons for this choice: mainly, the learning model on which it was designed (constructivism); the fact that it conformed to formalized educational software design criteria (Marquès, 1995); and the fact of being a monothematic web (skills, techniques and strategies for learning and studying), which meant that the test subjects would have different levels of knowledge of the content.

The next step was to introduce the four subjects to the Aprendre a aprendre web site and ask them to resolve a specific task on note-taking which would require extracting information from the site.

At first, we confined ourselves to observing and recording, by means of a computer program, the route followed by each of the test subjects. The program recorded each step the users took and the time taken. Observations were conducted while the subjects were carrying out the task, and at the same time an audio recording was made of everything they said. On a record sheet we noted down any incidents or relevant facts for later analysis. Afterwards we handed out a short questionnaire to find out what procedures the subjects thought they had followed in resolving the task. They were asked to draw a diagram of the structure of the site to find out how capable each was of perceiving the route she had taken in resolving the task, her reasons for adopting this approach, and how aware she was of having done so.

Once the field work and the empirical analysis of the information we had gathered was complete, we focussed on identifying the most significant aspects of the test subjects' approach to the task they were asked to complete.

Subject A: (Es/Ec) Planned in advance the steps she would take, had a clear idea of her objective and adapted her search to achieve it. She ignored any move that did not bring her closer to her goal. Knowing how to navigate and knowing the subject, she did not read the information she found in depth but skimmed through it. She selected the appropriate links and reached the goal by a short route. Because she knew the subject she was able to anticipate the conceptual framework, even though one of the points did not tie in with her own conception (error when she clicked on tools rather than on strategies). Her diagram of the structure of the web site was a reworking of her own conceptual framework, with the new knowledge built into it. She had a very clear idea of the blocks of content and how they were linked.

Subject B: (Ns/Ec) Planned in advance the steps she would take, gradually gained a clearer idea of the objective, and adapted her search to achieve it. She ignored any move that did not bring her closer to the goal. She knew how to navigate but she did not have a prior concept of the subject, and so she read each block of information at length. She selected the appropriate links and reached the goal by a short route, though it took her much longer than Subject A.

Because she was not familiar with the subject, she took a more linear route to the required information. She paid attention to the site plan. Her diagram of the structure of the web site was basically a navigational map. She had a very clear idea of the blocks and how they were linked, but she could not anticipate their content.

Subject C: (Es/Nc) Did not plan in advance the steps she would take, but she did have a clear idea of the objective. Because the web environment was so new to her she spent some time looking at and interacting with the links. After that, she ignored any move that did not bring her closer to the goal. She did not know how to navigate, so she made some mistakes. Because she knew the subject, once she had found the required block of content, her style of reading changed and became faster. She selected the appropriate links and reached the goal by a longer route than Subjects A or B, although she did not take as long as Subject B because she did not spend so much time reading once she arrived at the required block. From there she skipped several levels until she found the point she was looking for. The design of the web site made her pay less attention to the subject matter.

Her diagram of the structure of the site was a reworking of her own conceptual framework, with the new knowledge on the content which she imagined in a completely linear rather than a branching pattern built into it.

Subject D: (Ns/Nc) Did not plan the search process in advance. She started by trial and error and made no attempt to monitor and control her progress until she got lost. She proceeded by clicking on menus and reading them, which took up quite a lot of time. Of the four test subjects she was the one who took longest. At the end she seemed incapable of imagining the structure of the web site and had only a vague idea of the most general blocks.

Everything would seem to suggest that before users are given access to hypertextual media it would be helpful if they acquired some search and navigation strategies to guide them. If they had had such strategies, the two novices would probably not have lost their way during their web search. Getting lost in this hypertextual space gives rise to cognitive conflict in the searcher; web sites should try to anticipate this conflict and offer some kind of response to help searchers redirect their search.

Also, hypertextual systems demand a large amount of reading. It is advisable, therefore, when designing large hypermedia documents, to take the prior knowledge of target users into account. This would influence not only the way learners acquire knowledge, but also the way they choose hypertext content, the way they organize their search, and the way they digest information. Experts in computers are also experts at reading texts in hypermedia format and so are able to grasp the "big picture", whereas novices in the use of computers are also novices when it comes to reading hypertextual documents, and this makes them pay more attention to secondary details.

In our study we also observed how the computer novices expected to find information organized in a linear manner, following the sequence of the subject matter. When they did not find any such linearity, it took them some time to gain control of their search activity. The computer experts, in contrast, already knew that the information would most likely not be organized in a linear manner, and this helped them guess how the information might be structured. The computer expert who was also an expert in the subject was the only one who succeeded in gaining an overview of the site structure as a whole. The others obtained only a fragmentary view. The computer novice who was an expert in the subject had a clear grasp of the concepts, but did not know how they might be linked, while the computer expert who was new to the subject had a better idea of how the concepts might be linked, but could not remember the content of each node she had seen, much less imagine the content of those she had not seen.

Hypertext designers should know and bear in mind the characteristics of the target users, their goals, and the strategies they tend to adopt. This means that it is not enough to provide technologically advanced tools, as we cannot be sure that the new information systems will be appropriate for learners. We must bear in mind that hypertext allows us to represent information in a great variety of symbolic languages. This is something that web designers should exploit; they should respect the cognitive preferences of the users by integrating a greater variety of symbolic systems in the same hypertext.

Conclusions

New information and communication technologies are gradually bringing about a revolution which is already beginning to be considered as far-reaching and significant as its predecessors. This, however, is a revolution based on information and knowledge, which are likely to become a resource that changes the way we work and the way we live.

We must stop thinking that there is a standard level of knowledge that all students must attain. Everything

seems to indicate that knowledge is constantly being renewed; and the knowledge that was sufficient yesterday is incomplete or limited today. We should not be so concerned about what the new generations know when they leave school or university, but whether or not they have the ability to seek and find the information they need when they need it. And there is no doubt that, from this perspective, teaching students to find information is one of the great challenges of the immediate future; in order to find the information they need, they need certain procedural and strategic knowledge that they often lack.

The WWW could (and should) be a tool that helps students to become more independent in their learning, by placing a world of information within their grasp. It offers ready prepared, structured information. For example, it allows us to start our research in a library in Canada, continue it in a library in Australia, and finish it off in a library in South Africa. It calls into question the idea of linear reading as the only way to approach a text, and invites us to reflect on the possibility of following personalized routes, according to the information required.

In our survey-based descriptive study we found that Internet users often lack search resources (what information to look for, where, and how), tend always to use the same tools regardless of the objective and purpose of the search, fail to adapt their search method to the search objective, and, on occasions, have difficulties with English as the most common language on the Internet.

The skills required for finding information on the Web can be said to make up a new category of informational procedures. We call them procedures because they consist of a range of techniques and skills that can be learned, and we describe them as informational because they confer the ability to seek, select, interpret and reuse information. Informational procedures are nothing new: they are inherent in study and intellectual work: consulting dictionaries, knowing how to find one's way around a library, extracting relevant information from a document, presenting well structured dossiers, etc.

For all the reasons just mentioned, Internet search strategies should be considered a new learning content in obligatory and post-obligatory education.

Continuing this same line of argument, in view of the last two of our studies we can say that the mental operations carried out by expert, strategic searchers are different from those carried out by non-strategic or novice searchers. Strategic searchers engage in a process of planning review and monitoring evaluation, whereas non-strategic searchers proceed on the basis of association reproduction and automatization.

The way we look for information on the Internet or in hypertextual documents varies considerably, depending on what we want the information for, how much we already know (about the subject and/or about Internet resources), how specific the information is, our interest, our attitude, the language in which the search is conducted, the time we have available, the characteristics of the computer medium, the time of day when we connect, etc. The person who conducts the search should be able to identify relevant variables or conditions and plan her actions accordingly. This implies knowing, managing and adapting one's own navigation process, and this is a skill that can and must be learned.

To be successful in a search, therefore, it is not just a question of knowing certain things about the search medium or the specific information content. It is equally important to have a grasp of the language in which we want to communicate, to be familiar with different search procedures, and above all to know how to use them appropriately in the context of the objectives and conditions of the search. Added to this we need the ability to decide which of the documents we find are most trustworthy and will be most useful to us but that is something we will have to leave for a later study.

Bibliography

Adell, J. & Bellver, C. 1995. *La Internet como telaraña: el World–Wide Web. Métodos de Información*, vol. 2, n. 3.

Barberà, E. & Monereo, C. 1998. *Disseny instruccional de les estratègies d'aprenentatge escenaris educatius no-formals*. In Monereo, C. (Ed.), *Estratègies d'aprenentatge. Assessorament i formació del professorat*, vol.1. Barcelona: Ediuoc.

Bartolomé, A. 1999. *Hipertextos, hipermedia y multimedia: configuración técnica, principios para su diseño y aplicaciones didácticas*. In Cabero, J. (Ed.), *Medios audiovisuales y nuevas tecnologías para la formación en el siglo XXI*. Murcia: Diego Marín.

Cebrián, J. L. 1999. *La red: cómo cambiarán nuestras vidas los nuevos medios de comunicación*. Madrid: Taurus Ediciones – Grupo Santillana.

Cornella, A. & Rucabado, J. 1996. *Les autopistes de la informació. Descripció i impacte*. Barcelona: Proa – Columna.

Marquès, P. 1995. *Software educativo: guía de uso y metodología de diseño*. Barcelona: Editorial Estel.

Martí, E. 1992. *Aprender con ordenadores en la escuela*. Barcelona: ICE. Universitat de Barcelona/Horsori.

Martínez, J. A., Ros, F. & Santillana, I. 1996. *Las autopistas de la información*. Madrid: Dominiós.

Parra, B. 1998. *Cómo buscar información en internet*. Madrid: Anaya Multimedia.

Vicent, A. 1996. *¿Cómo buscar?*. European Schoolnet [WWW document]. URL <http://wfs.vub.ac.be/schools/timeline/search/buscar>.

Vizcarro, C. & León, J. 1998. *Nuevas Tecnologías para el aprendizaje*. Madrid: Pirámide.