Information Searcher-Provider Fit through Information Presentation and Visualization Techniques

Naresh Kumar Agarwal[a1] and Danny C. C. Poo[a]
[a] *School of Computing, National University of Singapore, 3 Science Drive 2, Singapore 117543*

Human-Computer Interaction (HCI) is all about the way in which people interact with computer systems. This study is concerned with the cognitive and visual aspects of HCI for Information Retrieval and its implications in helping a user find answers to his questions on the web. While the Internet has developed rapidly during the last 15 years, search engines have concentrated on providing a one-size-fits-all model for users. Users today need answers (and ways to get them fast) instead of simply being presented with a series of links. We suggest visualization and presentation techniques to help achieve a fit between different searcher modes and the services provided by a search service provider. Such an Information Searcher-Provider Fit will lead to effective search and a satisfied searcher.

Keywords: Human-Computer Interaction for Information Retrieval, Information Retrieval, Information Visualization, Web, Cognitive, Search, Information Searcher, Information Provider, Searcher Modes, Profile, Presentation, Fit, Adaptation.

1 INTRODUCTION

The Internet and the World Wide Web have developed rapidly during the past 15 years. It is being used increasingly by a large number of people across different countries and cultures for a variety of reasons, including browsing, shopping, research, education, collaboration, data sharing, online banking, etc. It has given rise to what is commonly termed as 'information overload'. The primary need now is to help different people with different cognitive levels and needs at different points in time find information they need at the right place, the right time and in the right format.

Effective user interactions with the vast amount of available information are essential to the success of the web in general and of the specific organizations, businesses and individuals that increasingly rely on it. Human-Computer Interaction (HCI) is all about the way in which people (individuals, groups or organizations) interact with computer systems [1]. It is concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them [2]. HCI involves *physical* aspects of the interaction (are the keys spaced right?), *perceptual* aspects (is the text color easy to see against the background?), *cognitive* aspects (will these menu names be understood?), and *social* aspects (will people trust each other on this auction site?) [1].

This study is concerned with the cognitive aspects of HCI for Information Retrieval and its implications on helping a user find answers to his questions on the Web. We also recommend providing presentation and visualization elements (physical/perceptual aspects of HCI) to help match the needs of the information searcher and the services provided by an information provider.

2 MOTIVATION

Cognitive science has been used in HCI to help understand how users interact with particular applications. However, understanding this in the context of Web is much more complex. A person's whole life – from early education to work to filing tax returns to shopping is increasingly being influenced or drawn into this global networked information structure. It is not just a technical artifact but part of the cultural backdrop of day-to-day life. This has far-reaching social implications and fundamental cognitive effects – we think differently because of the Web [1].

We often photocopy articles as a surrogate for reading them or have a sense of accomplishment after an Internet search as we download, but do not read, PDF files. In the post-Internet society, rather than knowing what we need to know, it is often more important to know how to find out what we need to know. As information becomes instantly globally available, this metacognitive knowledge, the about-information information, becomes increasingly important and it is not yet clear how this will change our cognitive engagement with the world [3]. As interface designers, we need to be aware of this because we design systems for this emerging cognitive demographic as well as design systems that shape it [1].

[1] Corresponding Author: Naresh Kumar Agarwal, School of Computing, National University of Singapore, 3 Science Drive 2, Singapore 117543, Email: naresh@comp.nus.edu.sg URL address: http://www.comp.nus.edu.sg/~naresh

The information available on the web has the following characteristics: a) the volume of information is huge; b) the type of information varies widely (e.g. from personal web pages to academic articles, from raw scientific data to online products and services); and c) information is not well structured. To retrieve information from the web, users often need to use web search engines to locate the websites containing information. Web search engines serve as catalogs of the Web. They index the web pages by using 'spiders' or 'robots', which crawl from site to site and create a database that stores indices of web pages on the Web [4].

However, finding relevant information on the Web is not easy. Unstructured information in large volume makes it difficult to efficiently index the Web pages (though this has been improving over the years). Search engines suffer from another major drawback – they make an underlying presumption that the user can formulate on-point queries to effectively narrow down the volume of information available [5]. Inefficient indexing and inaccurate search queries could easily result in millions of hits for a single search query [6]. Also, the interests of the users vary with time and cannot be represented by a fixed set [5] or a 'one-size-fits-all' model widely prevalent in the search engines of today.

The emphasis should hence be on addressing questions posed by users, through facilitating information search and knowledge discovery [7][8]. To address this, information providers should attempt to answer/address actual questions posed by the searcher rather than present them with a series of links. *'People are very impressed with web searches today but it's really quite poor compared to what it should be…a bunch of links that sort of start a treasure hunt that on average takes about 11 minutes'* (Bill Gates, Live! Forum, Singapore 1 July 2005).

Thus, there is a need for fit between the facilities provided by information providers (such as search engines) and the needs of information searchers.

3 NEED FOR FIT

Inspired by the Task-Technology Fit Framework of Daft & Lengel [9], we propose a research model for Information Searcher-Provider Fit (see Fig. 1). According to the model, in order to achieve effective search and to ensure searcher satisfaction, there must be a 'Fit' between the exact needs of the Information Searcher and the services provided by the Information Provider.

This would, of course, be the ideal case and is easier said than done. The establishment of this 'Fit' requires a departure from the current 'one-size-fits-all' model and the customization of search results based on the specific needs (and an understanding of the knowledge level in the area of search) of different searcher modes (discussed in Section 4).

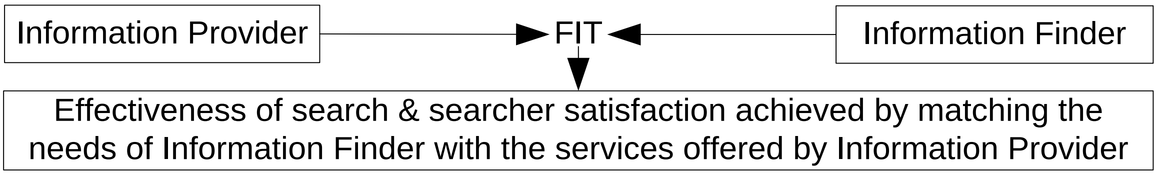

Fig. 1. Research Model. In order to achieve effective search and to ensure searcher satisfaction, there must be a 'Fit' between the exact needs of the Information Searcher and the services provided by the Information Provider.

Fit or adaptation will help in inching toward the goal of universal access to websites by tailoring the interface to the perceptual and cognitive capabilities of various user groups [10]. People may have a hard time stating why they are doing the actions that they are doing in an information seeking task. The search service provider must provide mechanisms for the searcher to determine and specify the mode he is in while carrying out the search. On getting and knowing the required context, the search engine must work interactively with the searcher until the searcher retrieves the search results relevant to his needs. This must happen in a manner and time frame which leaves the searcher satisfied.

4 AT THE HEART – THE USER

At the center of human-computer interaction is the user. In fact, it is said that many techniques and methods used in HCI succeed only insofar as they focus the designer on the user. Good designers get to understand their users by watching them, talking to them and looking at the things they produce. One technique that is used to help build this user focus is to produce profiles or *personae* of expected users. A design team may decide on several personae early in the design process typical of different user groups [1].

Much research has been carried out investigating users' abilities to navigate through the web to find desired information (e.g. [11], [12], [13], [14]). These studies highlight users' varied difficulties in navigating hypertext/hypermedia.

Web interfaces for search should accommodate not just users with different perceptual-motor capabilities and skills, but also the needs of the same person at different points in time when searching for different pieces of information.

Knowledge discovery can happen if information providers understand the modes searchers are in at different points of time, and then employ the best mechanism to serve them. This will lead to a fit or adaptation, as highlighted in Section 3. Adaptation is a multi-faceted process that can be analyzed across three main axes, namely the *source* of adaptation knowledge (knowledge available at start-up, knowledge derived at runtime), the *level* of interaction at which it is applied (semantic, syntactic and lexical), and the *type* of information on which it is based [15]. This study focuses on the source of adaptation knowledge.

Search and classification results must satisfy 4 basic categories of users [16]. We term these 4 categories of users (searcher modes) as a) novice b) data gatherer c) known-item searcher and d) focused searcher. Depending on the context of data one is searching for and the domain knowledge the person has in the field of search, the same person may be in one of the four[2] modes [17] (see Table 1).

Table1. Four modes/activities of Information Searchers. Goal-oriented search must satisfy four basic categories of searchers [16].

Searcher mode during a particular search	Searcher need during a particular search	Prior Domain Knowledge	Requirement of Specific Information
Novice	Needs information about a topic he is not familiar with in preparation for starting a new project	0 (no)	0 (not yet)
Data Gatherer	Needs information about a topic he *is* knowledgeable about and is therefore in data-gathering mode	1 (yes)	1 (yes)
Known-item Searcher	Has a good idea what he is looking for, knows that a given document or piece of data exists; simply needs to locate it	X (don't care)	1 (yes)
Focused Searcher	Needs a very specific answer to a specific question	X (don't care)	1 (yes)
Surfer[2]	Doesn't need anything in particular; searching purely for entertainment	X (don't care)	0 (no)

Having identified the searcher modes, the need now is for techniques and mechanisms that information/search service providers can employ to help achieve a 'fit' for each searcher mode.

5 INFORMATION VISUALIZATION TECHNIQUES

The power of the Web is linkage – users are free to follow their own interests, clicking on links as they wish. The downside is that after a period of browsing, the user can feel utterly lost in this hyperspace. There are two aspects to getting lost: 1) *Content*. The Web encourages a contextless style – like giving children an encyclopedia at the first day of school and saying 'learn'. The designer cannot know what visitors will have seen before or where they have come from. Usage of breadcrumbs, headers and menus helps provide a weak semblance of a shared context [1]. 2) *Spatial disorientation*. Miller's [18] classic paper showed that we have a relatively small working memory: 7 ± 2 chunks of information. This is all we have to keep track of where we have been. Websites employ site maps, breadcrumbs or other techniques to help users know where they are. Between sites, we have to rely on the browser's back button and history [1].

To rescue the user lost in hyperspace, an information searcher must be provided with appropriate visualization and presentation techniques based on the searcher mode he is in. A research report by Vividence [19] noted that the most common usability problem among 69 websites was poorly organized search results.

[2] The surfer is not looking for anything in particular and is just entertaining himself, so we don't count him in our typology of searcher modes.

Appropriate visualization techniques can enable the user to gain an overview of the entire information item collection, zoom in on items of interest, and filter out uninteresting items, get details of single or multiple items on demand, view the relationships among items or keep a history of actions to support undo, replay and progressive refinement [20]. We propose 3 techniques that can help meet the needs of searchers in different modes.

Taxonomy, a structured way to categorize information, provides a subject-based classification that arranges the terms in a controlled vocabulary into a hierarchy. It enables navigable search results that help users browse rather than search. Humans can rapidly navigate taxonomies to find high concentrations of topic-specific, related information [16]. Chen and Dumais [21] developed and studied an interface that organized search results into hierarchical categories. They found that participants in their study preferred the categorized results page over the typical list-style results page. They were also able to find information 50% faster than when using the list style of results.

A taxonomy should be highly useful to the novice who is just starting to work on a new project and does not know exactly what he should be searching for. A known-item searcher who is simply trying to locate a document or piece of information will be able to browse iteratively through the nodes of a taxonomy or different branches until he is able to locate the item of search. The data gatherer and the focused searcher have prior domain knowledge in the area of search and should be able to use a taxonomy and the search feature interchangeably.

Multiple Taxonomy Views. While taxonomies are flexible structures and can be developed to cover many different topics to any desired level of granularity, dynamic classifications are even more powerful. These allow search results to be organized in real-time into classification views that are selected by the user in order to view information from various perspectives [16]. Similarly, faceted classifications work by identifying a number of facets (different axes along which documents can be classified) into which the terms are divided e.g. classifying by color, geography, subject, etc. Facetmap.com uses a demo to classify wines into different facets - variety, region and price.

Multiple Taxonomy Views would be highly useful to the novice in helping him choose the view of choice. The different views might also benefit the known-item searcher in locating a document faster, since there would be different paths to arrive at the same document or piece of data. The focused searcher will be able to quickly arrive at the answer to his question utilizing a suitable view. The data gatherer would also benefit from multiple views though he might tend to use a particular view more than others – one that more closely fits in with his prior domain knowledge.

Concept maps. Concept maps are meaningful diagrams where concepts are represented by words enclosed within geometric figures, with relations among concepts expressed by words that label a directional, connecting line. Search results could also be organized as concept maps. Studies have shown that the appropriate use of concept maps for navigation can help people find topics more easily, provide easier, less frustrating access to information and can also result in learning gain [22].

Concept maps would have the greatest fit for the novice in being able to get an overall conceptual view. They would be useful to the known-item searcher in locating documents. A data gatherer is likely to concentrate on certain nodes of the map. A focused searcher will be able to zoom into a node of the map to get an answer to his question.

6 RESEARCH METHOD

Experiments will be carried out to determine what combination of visualization techniques (taxonomy, multiple taxonomy views, concept maps) is best suited for each searcher mode (novice, data gatherer, known-item searcher and focused searcher). Independent variables will be visualization technique and searcher mode. Effectiveness of search and searcher satisfaction will be measured to determine the degree of fit.

7 CONCLUSIONS AND IMPLICATIONS

This paper highlights that the currently prevalent ‘one-size-fits-all’ model for search is inadequate to serve the needs of different users at different points in time. In order to achieve a fit between the needs of the searcher and the mechanisms provided by the information provider, the user’s prior domain knowledge and requirement of specific information must be taken into account. Simply providing a set of links for the user to follow may not be enough anymore. We suggest appropriate visualization and presentation techniques to help achieve this fit. The right combination of techniques suited for each searcher mode will be determined empirically through experimentation. Apart from visualization techniques, future research will also include studying the impact of other mechanisms (such as personalization, contextual cues, semantics, local search, specialty search, as well as currently available general-purpose search engines) on different searcher modes to help achieve Information Searcher-Provider Fit.

REFERENCES

[1] Dix, A. Human-Computer Interaction and Web Design. 2005. In R. W. Proctor & K. -P. L. Vu (Eds.) Handbook of Human Factors in Web Design, Mahwah, New Jersey: Lawrence Erlbaum, pp. 28-47.

[2] Hewett, T. T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G., & Verplank, W. 1992. Human-Computer Interaction. In ACM SIGCHI Curricula for Human-Computer Interaction, New York: ACM, pp. 5-28 Available at http://sigchi.org/cdg [Accessed 16 June 2006]

[3] Dix, A., Howes, A. and Payne, S. 2003. Post-web cognition: evolving knowledge strategies for global information environments. International Journal of Web Engineering Technology, 1, pp. 112-126. Available from www.hcibook.com/alan/papers/post-web-cog-2003/ [Accessed 16 June 2006]

[4] Fang, X., Chen, P., & Chen, B. 2005. User search strategies and search engine interface design. In R. W. Proctor & K. -P. L. Vu (Eds.) Handbook of Human Factors in Web Design, Mahwah, New Jersey: Lawrence Erlbaum, pp. 193-210.

[5] Narayanan, S., Koppaka, L., Edala, N., Loritz, D., & Daley, R. 2004. Adaptive Interface for Personalizing Information Seeking. CyberPsychology & Behavior, 7(6), pp. 683-688.

[6] Fang, X., & Salvendy, G. 2000. Keyword comparison: A user-centered feature for improving web search tools. International Journal of Human Computer Studies, 52(5), pp. 915-931.

[7] Marchionini, G. 1997. Information seeking in electronic environments. Cambridge Series on Human-Computer Interaction, UK: Cambridge University Press.

[8] Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (eds.) 1996. Advances in Knowledge Discovery and Data Mining, Menlo Park, CA, The AAAI Press.

[9] Daft, R., & Lengel R. 1986. Organizational information requirements, media richness, and structural design. Management Science, 32(5), pp. 554-571.

[10] Proctor, R. W., & Vu, K -P. L. 2004. Human factors and ergonomics for the Internet. In H. Bidgoli (Ed.), The Internet encyclopedia. New York: Wiley, 2, pp. 141-149.

[11] Calvi, L. 1997. Navigation and disorientation: A case study. Journal of Educational Multimedia and Hypermedia, 6, pp. 305-320.

[12] Edwards, D. M., & Hardman, L. 1989. "Lost in hyperspace": Cognitive mapping and navigation in hypertext environment. In R. McAleese (Ed.), Hypertext: Theory into practice, Oxford, England: Intellect, pp. 105-125.

[13] Kim, H., & Hirtle, S. C. 1995. Spatial metaphors and disorientation in hypertext browsing. Behavior and Information Technology, 14, pp. 239-250.

[14] Stanton, N. A., & Baber, C. 1994. The myth of navigating in hypertext: How a "bandwagon" lost its course! Journal of Educational Multimedia and Hypermedia, 3, pp. 235-249.

[15] Stephanidis, C. 2001. The concept of unified user interfaces. In C. Stephanidis (Ed.), User interfaces for all – Concepts, methods and tools. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 371-388.

[16] Papadopoullos, A. 2004. Answering the Right Questions about Search. EContent Leadership Series - Strategies for...Search, Taxonomy & Classification, Supplement to July/August EContent and Information Today, pp. S6-S7. Available at http://www.procom-strasser.com/docs/Convera_Right_Questions.pdf [Accessed 16 June 2006].

[17] Agarwal, N.K., & Poo, D.C.C. 2006. Meeting knowledge management challenges through effective search. *Int. J. Business Information Systems*, 1(3), pp. 292-309.

[18] Miller, G. 1956. The magical number seven, plus or minus two: some limits on our capacity to process information. Psychological Review, 63(2), pp. 81-97.

[19] King, M. A. 2001. Tangled Web, Special Report, Vividence, Etailer's Digest, August 1, Available from http://etailersdigest.com/resources/Specials/Tangled_Web.htm [Accessed 16 June 2006]

[20] Klusch, M., Andre, E., Rist, T., Ishida, T., & Nakanishi, H. 2005. Interactive information agents and interfaces. In R. W. Proctor & K. -P. L. Vu (Eds.) Handbook of Human Factors in Web Design, Mahwah, New Jersey: Lawrence Erlbaum, pp. 211-236.

[21] Chen, H., & Dumais, S., "Bringing order to the web: Automatically categorizing search results," Proc. Conference on Human Factors in Computing Systems, CHI 2000, pp. 145-152, April 2000.

[22] Hoffman, R. R., Coffey, J. W., Novak, J. D., & Canas, A. J. 2005. Applications of concept maps to web design and web work. In R. W. Proctor & K. -P. L. Vu (Eds.) Handbook of Human Factors in Web Design, Mahwah, New Jersey: Lawrence Erlbaum, pp. 156-175.