Searchers' Criteria For Assessing Web Pages

Anastasios Tombros

Department of Computing Science, University of Glasgow, Glasgow, U.K.

tombrosa@dcs.gla.ac.uk

Ian Ruthven

Department of Computer and Information Sciences, University of Strathclyde, Glasgow, U.K.

Ian.Ruthven@cis.strath.ac.uk

Joemon M. Jose

Department of Computing Science, University of Glasgow, Glasgow, U.K.

jj@dcs.gla.ac.uk

ABSTRACT
We investigate the criteria used by online searchers when assessing the relevance of web pages to information-seeking tasks. Twenty four searchers were given three tasks each, and indicated the features of web pages which they employed when deciding about the usefulness of the pages. These tasks were presented within the context of a simulated work-task situation. The results of this study provide a set of criteria used by searchers to decide about the utility of web pages. Such criteria have implications for the design of systems that use or recommend web pages, as well as to authors of web pages.

Categories and Subject Descriptors
H.3.3 [**Information Search and Retrieval**]: – *search process*
H.5.4 [**Hypertext/Hypermedia**]: – *user issues*

General Terms
Design, Human Factors

Keywords
World Wide Web, user studies, evaluation, empirical studies

1. INTRODUCTION
Studies on peoples' perceptions of the relevance of information demonstrate that a range of factors affect human judgements of relevance. Barry [1], for example, studied factors that influenced searchers' criteria for relevance when reading textual documents. However, such studies often only consider formal textual documents, such as journal and conference articles, rather than the wide range of formally and informally-produced multimedia documents found on the web.

The motivation behind this study has been to gain a better understanding of what features make a web page useful for information seeking. We observe the decisions made by web users whilst searching on given information-seeking tasks. We gather, through think-aloud, system logging, and informal discussion, information on the relative utility of structural content (e.g. page layout, link structure), information content (e.g. use of text and multimedia) and aspects of quality (e.g. source of page or recency of information).

Our study has important implications for the design of systems that recommend or use web pages. A good example is web page

summarisation systems, e.g. [3]. Although such summaries have been shown to be useful for interactive searching, they generally only consider the textual content of the page. By utilising textual *and* non-textual aspects of pages, summarisation systems can provide a more complete picture of web pages.

2. METHODOLOGY
Twenty four searchers participated in the study. Each was given three search tasks to complete. The searchers were asked to discuss, in the form of think-alouds, their perceptions of what features helped them decide on the usefulness (or non-usefulness) of the web pages they chose to view. There were no restrictions imposed on the search engines or web sites searchers could visit. User desktop actions (e.g. web pages accessed, mouse movements, etc.) and speech at the duration of each task were recorded in the form of video sessions using the *Camtasia* software package (*Techsmith Corporation*, *www.techsmith.com*). The sessions were subsequently analysed by one of the experimenters.

2.1 Search Tasks
Each of the subjects was asked to complete three search tasks. The tasks were placed in the context of a simulated work task situation [2]. In this way, users are encouraged to treat the information seeking task as a personal task, searching as though the task was their own. The three tasks also simulated different types of searches.

In the first task, a *background search*, searchers were asked to find as much general background information on the demographics of the Internet as possible. For the second task subjects were asked to decide on the best hi-fi speakers available in their own price range (*decision task*). The third task, a *many items search*, asked the subjects to compile a list of interesting things to do over a weekend stay in the city of Kyoto. The order of presentation of the tasks was rotated across subjects.

2.2 Analysing User Sessions
We analysed each recorded user session, and extracted web page features that were mentioned by searchers as indicating a page's usefulness (*positive mention*) or non-usefulness (*negative mention*) to a task (i.e. features were discovered through searchers' mentions rather than predefined by the experimenters). We thus gathered data for features in two lists, depending on the type of pages (useful or not) in which they appeared. After the processing of all user sessions, features were grouped under broader feature categories. In total, 24 features were grouped under 5 feature categories (Table 1).

3. RESULTS

In Table 1 we present the categories (left column) of features, and their corresponding features (right column). In total, we obtained assessments for 862 web pages (346 useful pages and 516 non-useful). The assessments were derived through 1,602 mentions of features (786 positive mentions and 816 negative). For 371 pages, a single feature was mentioned by searchers as indicating utility (78 cases) or non-utility (293 cases).

Table 1. Features and categories of features for web pages

Text	Content, Numbers, Titles/Headings, Query Terms, Text Quantity
Structure	Layout, Links, Links Quality, Table Layout
Quality	Scope/Depth, Authority/Source, Recency, General Quality, Content Novelty
Non-textual	Pictures
Physical Properties	Page Not Found, Page Location, Page Already Seen, Others

In Figure 1 we present the percentage of mentions of the most important features. Percentages in this figure have been calculated by considering both positive and negative mentions of features, and by considering mentions across all tasks. The overall relative importance of feature categories is that of Table 1 (Text with 44.7% of all mentions, Structure 21.6%, Quality 17.7%, Non-textual Items 8.9%, Physical Properties 7.1%). The two most commonly mentioned features (the general *content* of pages and textual parts which contain *numbers*) both belong to the *Text* category. Features from the second most important category (*Structure*), i.e. layout and links, also appear in the top places of the mentioned features, as well as the only non-textual item met at the study (pictures). The authority/source of the pages and the scope and depth of the pages' information (*Quality* category) also appear ranked high in Figure 1.

Features with a high tendency of being mentioned as single indicators of page (non-)utility can be potentially useful at helping users identify (non-)useful web pages. The content of web pages is the feature which is most often used as a sole indication of a page's utility, with 29% (132 mentions) of its total mentions made in this way. Textual parts of the pages which contain query terms are also high in this category with 38% (24 mentions) of their mentions made as single mentions. Other features that are commonly used as single indicators of utility and non-utility are the quality of the links, recency of information, scope and depth of the information, and the links in a web page.

Features which exhibit the tendency to mainly indicate either useful or non-useful pages are equally important, since they both inform searchers whether a certain page is worth their time and effort. Content remains the most commonly mentioned feature in both cases, with little variation between positive and negative mentions. However, links, authority/source, numbers and pictures have mainly positive mentions (approximately 70% of their total mentions are positive). On the other hand, links quality, general quality, scope/depth and layout have significantly more negative mentions (88%, 76%, 68% and 62% respectively). All features of the *Physical Properties* category have significantly more negative mentions, mainly due to their nature.
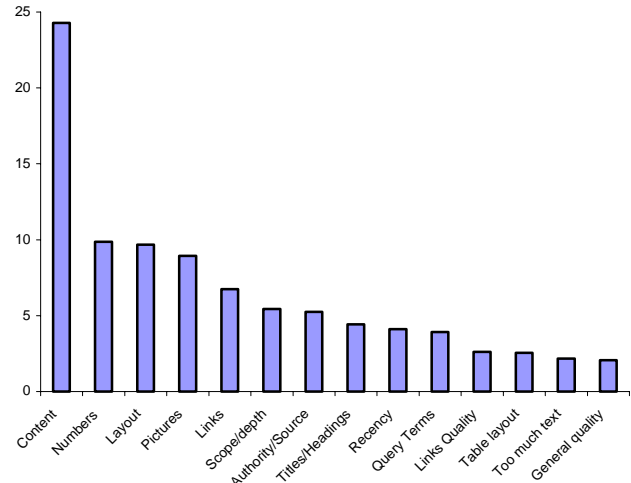
Figure 1. Percentage of mention of page features

Our results also suggest that the type of task affects the features that users employ, that certain features tend to co-occur and act as strong indicators of utility or non-utility, and that different features are employed across different stages of the search tasks by subjects. We plan to report on these results in detail in the future.

In conclusion, the results reported here confirm that various textual aspects of web pages (general content, textual parts containing query terms and numbers, text in the title and headings of pages), are important for determining the utility of pages to tasks. This is in agreement with findings regarding the criteria searchers employ when assessing textual documents [1]. However, structural and qualitative aspects of web pages (e.g. layout, links, scope/depth, authority/source) are also important, and should be considered. These results, along with the effect of task type and task stage on the features mentioned by searchers, have wider implications for the design of systems that use or recommend web pages.

4. ACKNOWLEDGEMENTS

5. REFERENCES

[1] Barry, C.L. Document representations and clues to document relevance. Journal of the American Society for Information Science, 49(14):1293-1303, 1998.

[2] Borlund, P. Experimental components for the evaluation of interactive information retrieval systems. Journal of Documentation, 56(1):71-90, 2000.

[3] White, R. W., Jose, J. M. and Ruthven, I. A task-oriented study on the influencing effects of query-biased summarisation in web searching. Information Processing and Management, 2003. in press.